

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 4, 2002, 08:38:05 : Search time 4179.27 Seconds
(without alignments)
6481.608 Million cell updates/sec

Title: US-09-052-089a-7
Perfect score: 2007
Sequence: 1 GTGGCGTGCAGCAAAATTG.....AAAAAAAAAAAAAAAAAAAA 2007

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 674847542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_estl:*
10: gb_estc:*
11: gb_hlc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1075.2	53.6	2696	11	AK012948	AK012948 Mus muscu
2	1067.2	53.2	1938	11	AK012786	AK012786 Mus muscu
3	894.4	44.6	959	9	AL560947	AL560947 AL560947
4	863.4	43.0	887	9	AL560912	AL560912 AL560912
5	777	38.7	1061	10	BM464844	BM464844 AGENCOURT
6	685.8	34.2	775	9	AL582241	AL582241 AL582241
7	659.8	32.9	769	10	BF038722	BF038722 601460641
8	654.4	32.6	702	10	BC764118	BC764118 602737137
9	651.4	32.5	888	10	BE869186	BE869186 601445180
10	651	32.4	784	10	BG120736	BG120736 602347389
11	650.2	32.4	752	10	BG682548	BG682548 602624362
12	632.6	31.5	908	10	BG420765	BG420765 602448558
13	632.4	31.5	668	10	BI758046	BI758046 603030577
14	629.8	31.4	645	9	AL559248	AL559248 AL559248
15	628	31.3	825	10	BE346959	BE346959 601071929
16	626.8	31.2	983	10	BE889971	BE889971 601512230
17	608.6	30.3	632	10	BF219707	BF219707 601296627

18	599.4	29.9	912	10	BG117979	BG117979 602351223
19	560.8	27.9	932	9	AL529568	AL529568 AL529568
20	556	27.7	619	9	AL580903	AL580903 AL580903
21	520.4	25.9	595	9	AL582265	AL582265 AL582265
22	498	24.8	702	10	BI151643	BI151643 602916189
23	490.4	24.4	905	9	AL563489	AL563489 AL563489
24	489.6	24.4	744	10	BF162255	BF162255 601770948
25	476.8	23.8	687	10	BF658170	BF658170 maag6b06.
26	462.8	23.1	523	9	AA564143	AA564143 nt44f02.s
27	461.2	23.0	547	9	BE031265	BE031265 129770 MA
28	443	22.1	571	10	BI341090	BI341090 36835 MA
29	422.6	21.1	495	10	BG235999	BG235999 na2f0f09.
30	410.4	20.4	478	9	AM248701	AM248701 2821007.3
31	406.2	20.2	519	10	BE304189	BE304189 601086421
32	398.6	19.9	554	9	AA684194	AA684194 vm68d10.s
33	387.4	19.3	589	10	BE334637	BE334637 us81g12.y
34	373.2	18.6	435	9	AI147333	AI147333 g963q01.s
35	365	18.2	455	9	AM207517	AM207517 UI-H-B11-
36	351.2	17.5	399	9	AI695169	AI695169 wa47h01.x
37	346	17.2	457	9	BE234038	BE234038 140628 MA
38	345.4	17.2	1015	10	BM462511	BM462511 AGENCOURT
39	334	16.6	379	9	AI633797	AI633797 tt28f05.x
40	334	16.6	379	9	AM002888	AM002888 wt02c11.x
41	331.8	16.5	380	10	BE552288	BE552288 hy05e12.x
42	328.4	16.4	394	9	AM207547	AM207547 UI-H-B11-
43	327.6	16.3	370	9	AM249133	AM249133 2821007.5
44	321	16.0	366	9	AI699987	AI699987 tt35f05.x

ALIGNMENTS

RESULT 1	AK012948	LOCUS	AK012948	2696 bp	mus musculus 10, 11 days embryo whole body cDNA, RIKEN full-length enriched library, clone:2810054N23.TRAF-Interacting protein, full insert sequence.
DEFINITION	AK012948	LOCUS	AK012948	2696 bp	mus musculus 10, 11 days embryo whole body cDNA, RIKEN full-length enriched library, clone:2810054N23.TRAF-Interacting protein, full insert sequence.
ACCESSION	AK012948	VERSION	AK012948.1	GI:12850018	
KEYWORDS	HTC; CAP trapper.	SOURCE	Mus musculus (strain:C57BL/6J) 10, 11 days embryo cDNA to mRNA, clone:2810054N23.		
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
AUTHORS	Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.				
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes				
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)				
MEDLINE	20499374				
PUBMED	11042159				
REFERENCE	3 (sites)				
AUTHORS	Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsuami,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsunoto,H., Sakauchi,S., Ikegami,T., Kashiwagi,K., Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watabiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,T., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.				
TITLE	RIKEN Integrated sequence analysis (RISA) system-384 format				
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)				

Dd	1560	-----AAAGTGTTTTGGAAGATGGCTTCCTTGACCA	1592
Oy	1643	GAGTAGGTGACACCAGACTGCTCTTCCTGCCTCACCCCTCCTAGAC	1702
Dd	1593	GTCCAAAGAGATGCCAAMAAACACTTCTGTGTTCACATCGCCCTGCAC--AAC	1650
Oy	1703	TGGAGCTGACATGACACCCACTGATCTGTGACAGAGTCTGCT--TCTGTGCGAGG	1761
Dd	1651	TGGGAGCCACATGACCATGTTACTGTTCCTCCATGACAGAGGCTACTTCAGATGTCAGG	1710
Oy	1762	TCTGTATTATGCCATGATCAGATGTGTGTCAGACTTCTTGCGGCTGGAACACAGCT	1821
Dd	1711	TTTTCTATTATGCTAACACACAGTGTGGCTGACATCTTTGTTTATAGAAGAGGTC	1770
Oy	1822	ACTT 1825	
Dd	1771	ACAT 1774	
RESULT	3		
AL560947			
LOCUS	ALS60947	959 bp	mRNA linear EST 16-FEB-2001
DEFINITION	ALS60947 LTL_NFL010_BC2 Homo sapiens cDNA clone CSODL005M09 5		
ACCESSION	ALS60947		
VERSION	ALS60947.1	GI:12907896	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	Li,W.B., Gruber,C., Jesses,J., and Polayes,D.		
JOURNAL	Full-length cDNA libraries and normalization		
COMMENT	Unpublished (2001)		
CONTACT:	Genoscope		
Genoscope - Centre National de Sequencage			
BP 191 91006 Evry cedex - France			
Email: seque@genoscope.cns.fr, Web : www.genoscope.cns.fr.			
FEATURES			
Source	1..959		
/organism="Homo sapiens"			
/db_xref="taxon:9606"			
/clone="CSODL005YM09"			
/clone_lib="LTL_NFL010_BC2"			
/sex="male"			
/tissue_type="B cells from Burkitt lymphoma"			
/note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA was primed with a NotI-Oligo(dt) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"			
BASE COUNT	259 a 237 c 275 g 185 t	3 others	
ORIGIN			
Query Match	44.6% Score 894.4 DB 9 Length 959:		
Best Local Similarity	97.6% Pred No. 5 Le-157:		
Matches 937: Conservative 3 Mismatches 14 Indels 6 Gaps 3			
Oy	20 GAAGCAAGCGGAGGCGGCG---CTCTACGAAGCGGAGCTGTAGACAGTTCTTTGGC	75	
Dd	1 GAAGCAAGCGGAGGCGGCGGCGGCGGCTTACGAAGCGGAGCTGTAGACAGTTCTTTGGC	60	
Oy	76 TGCCTGGGCCCCCTTGAATCAGCATCATGCTATCCGTGCTGTGACACTATCTCTCC	135	
Dd	61 TGCCTGGGCCCCCTTGAATCAGCATCATGCTATCCGTGCTGTGACACTATCTCTCC	120	

QY	136	GAC	TTCTTGATCACTCCCGGAGCTGGCGGCATCCACTGGGCCACACCTTCCACTTG	195
Db	121	GAC	TTCTTGATCACTCCCGGAGCTGGCGGCATCCACTGGGCCACACCTTCCACTTG	180
QY	136	CAG	TCCCTAATTTCAGTCCCTTTGAGACACACCACTGCGGACCTGCCAATGCGCGAATC	255
Db	181	CAG	TCCCTAATTTCAGTGGTTTTGAGCACACACCACTGCGGACCTGCCAATGCGCGAATC	240
QY	256	CAG	TTGGCAAAAGAACCATATCAATTAAGCTCTTCTTGATCTCTGGCCAGAGAGAG	315
Db	241	CAG	TTGGCAAAAGAACCATATATCAATTAAGCTCTTCTTGATCTTGGCCAGAGAGAG	300
QY	316	AAT	GTCTTGATCGAATATCTTAAAGAAATGAACCTGACCAATGTGAGAGCCAGCTTTCC	375
Db	301	AAT	GTCTTGATCGAATATCTTAAAGAAATGAACCTGACCAATGTGAGAGCCAGCTTTCC	360
QY	376	CAG	AAAGACAAAGAGAGAAACAGACAGCCAGCTCATCATGACACTCTCGGGGATACGCTG	435
Db	361	CAG	AAAGACAAAGAGAGAAACAGACAGCCAGCTCATCATGACACTCTCGGGGATACGCTG	420
QY	436	GAA	GAAAGCAATAGTCTACTGTGGTATCTCTGCAGCAGGCGCTTGGGCAAGCCGAGATGCTG	495
Db	421	GAA	GAAAGCAATAGTCTACTGTGGTATCTCTGCAGCAGGCGCTTGGGCAAGCCGAGATGCTG	480
QY	496	TGC	CTCCACACTGAAAAAACACAGATGAAGTACTTGAAGCAGACAGCAGATGAGACCAACAA	555
Db	481	TGC	CTCCACACTGAAAAAACAGATGAAGTACTTGAAGCAGACAGCAGATGAGACCAACAA	540
QY	556	GC	ACAAGAGAGAGCGGGCGCGCTCAGAGCAAGATGAAGACCATATGAGCAGATTGAGCTT	615
Db	541	GC	ACAAGAGAGAGCGGGCGCGCTCAGAGCAAGATGAAGACCATATGAGCAGATTGAGCTT	600
QY	616	CTA	CTCCAGAGCAGCTCCCTGAGGTGGAGAGGATGATCCCGACATGGTGTGGGACAG	675
Db	601	CTA	CTCCAGAGCAGCTCCCTGAGGTGGAGAGGATGATCCCGACATGGTGTGGGACAG	660
QY	676	TC	AGGGTGGAGACAGCTGGCTGTACTGTGTCTCTCAAGAAAGATAGAGAACTCA	735
Db	661	TC	AGGGTGGAGACAGCTGGCTGTACTGTGTCTCTCAAGAAAGATAGAGAACTCA	720
QY	736	AA	AGAGGACGAGAGCCCTCAGGGAGAGTGGCTGCAACAGCTGAGAGAGATTGTGTTCC	795
Db	721	AA	AGAGGACGAGAGCCCTCAGGGAGAGTGGCTGAC-ARCTGAGGAAAGATTGTGTTCC	779
QY	796	TCC	AGAGCAATGTCAGACAGTCTAGTCTGAAATGGAAATCGAGCCCAAGTTGAATGAG	855
Db	780	TCC	AGAGCAATGTCAGACAGTCTAGTCTGAAATGGAAATCGAGCCCAAGTTGAATGAG	839
QY	856	TC	AGCCCAAGAGACTTACAGAGTCTCTACAAGAAATCATGAGCCTGAAAAAGACCTA	915
Db	840	TC	AGCCCAAGAGACTTACAGAGTCTCTAMAAGGAATCATGAGCCTTGAAGAAGACCTA	899
QY	916	AC	GATGCTGTCAGAGAAACCTT-GAACCTGGCACAGTGGCCAGTGAAGCTGTGACCGCTT	974
Db	900	AC	GATGCTGTCAGAGAAACCTTGGAACTGTGCACAAATGGGSCAGTGAAGCTGTGACCGCTT	959
RESULT 4				
AL560912			887 bp	mRNA linear EST 16-FEB-2001
LOCUS	AL560912	LTI_NFL010_BC2	Homo sapiens	cdna clone CS0DL005YI08 5
DEFINITION	prime, mRNA sequence.			
ACCESSION	AL560912			
VERSION	AL560912.1	GI:12907828		
KEYWORDS	EST.			
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
AUTHORS	Li, W.B., Gruber, C., Jesssee, J. and Polayes, D.			
TITLE	Full-length cDNA libraries and normalization			

JOURNAL Unpublished (2001)
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, web : www.genoscope.cns.fr.

FEATURES
 Source
 1. .887
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="CS0DL005Y108"
 /clone_lib="LTI_NFL010_BC2"
 /sex="male"
 /tissue_type="B cells from Burkitt lymphoma"
 /note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com ;

BASE COUNT 251 a 219 c 245 g 171 t 1 others

ORIGIN

Query Match 43.0%; Score 863.4; DB 9; Length 887;
 Best Local Similarity 99.0%; Pred. No. 3.1e-151;
 Matches 878; Conservative 1; Mismatches 7; Indels 1; Gaps 1;

```

OY 84 CCCCTTGAGTCCAGCATCATGCTATCCGCTCTGTGCTACTATCTGCTCGACTTCTT 143
Db 1 CCCCTTGAGTCCAGCATCATGCTATCCGCTCTGTGCTACTATCTGCTCGACTTCTT 60
OY 144 CGATCACTCCCGGACGTCGCGCCATCTGCGGCCACACCTTCCACTTGGACGTCCT 203
Db 61 CGATCACTCCCGGACGTCGCGCCATCTGCGGCCACACCTTCCACTTGGACGTCCT 120
OY 204 AATTCAGTCTTTGAGACAGCAGCAAGTCGAGCTGGCCACAGTCCGAAATCCAGGTTGG 263
Db 121 AATTCAGTCTTTGAGACAGCAGCAAGTCGAGCTGGCCACAGTCCGAAATCCAGGTTGG 180
OY 264 CAAAAGAACCATTTATCATTAAGCTCTTCTTGTATCTTGCAGAGAGAGAGAAATGTCTT 323
Db 181 CAAAAGAACCATTTATCATTAAGCTCTTCTTGTATCTTGCAGAGAGAGAGAAATGTCTT 240
OY 324 GGATCGAATTTCTTAAAGATGAATGACATGTCAGAGCCAGCTTTCCAGAAAGA 383
Db 241 GGATCGAATTTCTTAAAGATGAATGACATGTCAGAGCCAGCTTTCCAGAAAGA 300
OY 384 CAAGGAGAAAGAGACAGCCAGTCATCATGACACTCTGGGGATAGCGGTGGAAGAG 443
Db 301 CAAGGAGAAAGAGACAGCCAGTCATCATGACACTCTGGGGATAGCGGTGGAAGAG 360
OY 444 CAATGCTACTGTGTATCTCTGACAGAGCCCTTGGGCAAGGCCAGATGTGTCTCTCAC 503
Db 361 CAATGCTACTGTGTATCTCTGACAGAGCCCTTGGGCAAGGCCAGATGTGTCTCTCAC 420
OY 504 ACTGAAAAAGAGATGAAGTACTTGAACAGCAGCAGAGATGAGCCAAACAGACAGAGA 563
Db 421 ACTGAAAAAGAGATGAAGTACTTGAACAGCAGCAGAGATGAGCCAAACAGACAGAGA 480
OY 564 GGAGGCGGGCGGCTCAGAGAGCAGATGAAGACCATGAGAGATGAGGCTTCTACTCCA 623
Db 481 GGAGGCGGGCGGCTCAGAGAGCAGATGAAGACCATGAGAGATGAGGCTTCTACTCCA 540
OY 624 GAGCAGACCTCCCTGAGGTGAGAGATGATCCAGACATGGGTGTGGACAGTACGGGT 683
Db 541 GAGCAGACCTCCCTGAGGTGAGAGATGATCCAGACATGGGTGTGGACAGTACGGGT 600
OY 684 GGAACAGCTGGCTGTACTGTGTCTCTCAAGAAAGAGTACGAGATCTTAAAGAGGC 743
  
```

Db 601 GGAACAGCTGGCTGTACTGTGTCTCTCAAGAAAGATAGCAGAAATCTAAAGAGGC 660

OY 744 ACGGAAGCCCTCAGGAGGAGGCTGACAGCTG -ACGAGAGATTGTCTTCCCGAAG 802

Db 661 ACGGAAGCCCTCAGGAGGAGGCTGACAGCTGAAAGAAAGAGATTGTCTTCCCGAAG 720

OY 803 GCAAGTTCAGACAGTCTACTCTGAATTTGATGATGAGCCAGCAAGTTAGAACTGACGCC 862

Db 721 GCAAGTTCAGACAGTCTACTCTGAATTTGATGATGAGCCAGCAAGTTAGAACTGACGCC 780

OY 863 AGAAGCACTTACAGAGTCTGACAGAGAAATCATGACCTGAAAAGACCTAACATGC 922

Db 781 AGAAGCACTTACAGAGTCTGACAGAGAAATCATGACCTGAAAAGACCTAACATGC 840

OY 923 TGCAGGAACCTTGAACCTGCCACCACTGGCCAGTGCAGTGCAGTGCAG 969

Db 841 TGCAGGAACCTTGAACCTGCCACCACTGGCCAGTGCAGTGCAGTGCAG 887

RESULT 5

LOCUS BM464844 1061 bp mRNA EST 05-FEB-2002

DEFINITION AGENCOURT_6428661 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:5504946

ACCESSION BM464844

VERSION BM464844

KEYWORDS BM464844.1 GI:18513886

SOURCE EST.

ORGANISM human.

REFERENCE Eukarya; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo. 1 (Bases 1 to 1061)

AUTHORS NIH-MGC http://mgc.nci.nih.gov/

TITLE National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)

JOURNAL Contact: Robert Strausberg, Ph.D.

COMMENT Email: cgabs-r@mail.nih.gov
 Tissue Procurement: Lou Staudt
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM12147 row: b column: 19
 High quality sequence start: 2
 High quality sequence stop: 641.

FEATURES
 Location/Qualifiers
 1. .1061
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5504946"
 /clone_lib="NIH_MGC_85"
 /tissue_type="lymphoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: Lymph; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.867 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."

BASE COUNT 282 a 272 c 297 g 202 t 8 others

ORIGIN

Query Match 38.7%; Score 777; DB 10; Length 1061;
 Best Local Similarity 97.2%; Pred. No. 3.9e-135;
 Matches 808; Conservative 0; Mismatches 21; Indels 2; Gaps 2;

```

OY 34 CGGGGGCTCTAGAGCGGAGCAGCTTCTTTGGCGCTGGCCCTTGAGT 93
Db 26 CGGGGGCTCTAGAGCGGAGCAGCTTCTTTGGCGCTGGCCCTTGAGT 85
OY 94 CCAGCATCATGCTATCCGCTGTGTGCACATATCTGCTCCGACTTCTTCATCACC 153
  
```

Db	86	CCAGCATTACGCGATCCGATGCTCTGTGACACTATCTGCTCCGACTTCTTGATCATTCC	145
OY	154	CGCGACGTGGCCGCCATCCACTGCGGCCACACCTTCCACTTGCAGTGCCTAATTCACTCC	213
Db	146	CGCGACGTGGCCGCCATCCACTGCGGCCACACCTTCCACTTGCAGTGCCTAATTCACTGG	205
OY	214	TTTGGACAGCACCAAGTCGGACCTGGCCACAGTGGCCAAATCCAGTTGGGCAAAAGAAC	273
Db	206	TTTGGACAGCACCAAGTCGGACCTGGCCACAGTGGCCAAATCCAGTTGGGCAAAAGAAC	265
OY	274	ATTATCAATTAAGCTCTTCTTGTGATCTTTCGCCAGAGAGAGAAATGCTTGGATGCAGAA	333
Db	266	ATTATCAATTAAGCTCTTCTTGTGATCTTTCGCCAGAGAGAGAAATGCTTGGATGCAGAA	325
OY	334	TTCTTTAAGGATGAACTGGACAAATGTCAAGGCCCTTTCCAGAAAACAGAGAGAA	393
Db	326	TTCTTTAAGGATGAACTGGACAAATGTCAAGGCCCTTTCCAGAAAACAGAGAGAA	385
OY	394	CGAGACACCAGATCATGACACTCTGGGGATACGCTGGAGAAAGCGCAATGCTCTCT	453
Db	386	CGAGACACCAGATCATGACACTCTGGGGATACGCTGGAGAAAGCGCAATGCTCTCTCT	445
OY	454	GTGCTATCTCTGCACAGAGCCTTGGGCAAGGCCAGATGCTGTGCTCCACACTGAAAAAG	513
Db	446	GTGCTATCTCTGCACAGAGCCTTGGGCAAGGCCAGATGCTGTGCTCCACACTGAAAAAG	505
OY	514	CAGATGAAGTACTTATGACACAGCAGAGATATGACAAACAAGACACAGAGAGCGGGCC	573
Db	506	CAGATGAAGTACTTATGACACAGCAGAGATATGACAAACAAGACACAGAGAGCGGGCC	565
OY	574	CGGCTCAGAGACCAAGATGAAGACCACTGAGAGCAGATTGAGCTCTCTCAGAGCGAGCTC	633
Db	566	CGGCTCAGAGACCAAGATGAAGACCACTGAGAGCAGATTGAGCTCTCTCAGAGCGAGCTC	625
OY	634	CCTGAGGTGGAGAGAGATATCCGAGACATGGGTGTGGACACTGACGCGTGGACAGCTG	693
Db	626	CCTGAGGTGGAGAGATATCCGAGACATGGGTGTGGAGACACTGACGCGTGGACAGCTG	685
OY	694	GCTGTACTGTGATCTCTCTCAAGAAAGATGCGAGATCTTAAAGAGGACAGGAAAGCC	753
Db	686	GCTGTACTGTGATCTCTCTCAAGAAAGATGCGAGATCTTAAAGAGGACAGGAAAGCC	745
OY	754	TCAGGGAAGTGGCTGACAGAGCTGAGGAAGGATTTGTTTCC -TCAGAGCAAGCACTTGCA	812
Db	746	TCAGGGAAGTGGCTGACAGAGCTGAGGAAGGATTTGTTTCC -TCAGAGCAAGCAAGTGA	805
OY	813	GACACTCTACTCTGAATGATCAGAGCCCAAG -TTTGAAGTGAAGTCAAGCC	862
Db	806	GACACTCTACTCTGAATGATCAGAGCCCAAGTTTGAAGTGAAGTCAAGCC	856

```

source
1. .775
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CSODL005Y108"
/clone_1db="LTI_NFL010_BC2"
/sex="male"
/tissue_type="B cells from Burkitt lymphoma"
/notice="Vector: pcwvSPORT 6; Site 1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pcwvSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifestech.com url : http://fulllength.invitrogen.com"
```

RESULT	6				
LOCUS	AL582241/c				
DEFINITION	AL582241	775 bp	mRNA	linear	EST 16-FEB-2001
ACCESSION	AL582241				
VERSION	AL582241				
KEYWORDS	AL582241.1	GI:12950030			
SOURCE	EST.				
ORGANISM	human.				
REFERENCE	Homo sapiens				
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
JOURNAL	1 (bases 1 to 775)				
COMMENT	L1,W.B., Gruber,C., Jesse,J. and Polayes,D.				
	Full-length cDNA libraries and normalization				
	Unpublished (2001)				
	Contact: Genoscope				
	Genoscope - Centre National de Sequencage				
	Bp 191 91006 EVRY cedex - France				
	Email: segr@genoscope.cns.fr, web : www.genoscope.cns.fr.				

Query Match	Similarity	34.2%	Score 685.8	DB 9	Length 775
Best Local	Similarity	95.2%	Pred. No. 4	Le-118	
Matches	738	Conservative	11	Mismatches	18
				Indels	8
				Gaps	4
QY	1045	GATCTCAATGCTACCTTTGATGTGATATCTCCCGACGCCGGCCCTCCAGCTCCAGCAT	1104		
Db	775	GATCTCAATGCTACCTTTGATGTGATATCTCCCGACGCCGGCCCTCCAGCTCCAGCAT	716		
QY	1105	GGTTACTACGAAAAAATTGGCTTAGAAGATCACACTCCCAATTCAGATGTCCCCAAG	1164		
Db	715	GGTTACTACGAAAAAATTGGCTTAGAAGATCACACTCCCAATTCAGATGTCCCCAAG	656		
QY	1165	AAGATATCGAAAGGCCCGACGAGAGAGTCCGAGCTCTCACTGGGAGGCCAGAGCTGTGA	1224		
Db	655	AAGATATCGAAAGGCCCGACGAGAGAGTCCGAGCTCTCACTGGGAGGCCAGAGCTGTGA	596		
QY	1225	GGAGAGCCAGATGAGAGAACTGGTTGGTGGCTTCCCTATTTTGTGCCGAATGCCATCTTA	1284		
Db	595	GGAGAGCCAGATGAGAGAACTGGTTGGTGGCTTCCCTATTTTGTGCCGAATGCCATCTTA	536		
QY	1285	GGCCGAAAAACAGCCAAAAAGGCCGAGGTACAGTCTCTTGGACGAAAGATGTGGTAAG	1344		
Db	535	GGCCGAAAAACAGCCAAAAAGGCCGAGGTACAGTCTCTTGGACGAAAGATGTGGTAAG	476		
QY	1345	ACAGAGCTTCGATGGGCTCGGTGGCCGGACAAAATTCATCCAGCTTACTGACACAGTCATG	1404		
Db	475	ACAGAGCTTCGATGGGCTCGGTGGCCGGACAAAATTCATCCAGCTTACTGACACAGTCATG	416		
QY	1405	ATCCGCCCATTTGCTGTTAAGGCCCAAGACCAAGGTTTAAGCAGAGGTTGAGGTTGAAGACC	1464		
Db	415	ATCCGCCCATTTGCTGTTAAGGCCCAAGACCAAGGTTTAAGCAGAGGTTGAGGTTGAAGACC	356		
QY	1465	GTGCTTCTCTTTCACAGGCCCAAGCTGGACACCTTCTGTGTCTGTGAACAAGTGAATC	1524		
Db	355	GTGCTTCTCTTTCACAGGCCCAAGCTGGACACCTTCTGTGTCTGTGAACAAGTGAATC	296		
QY	1525	TGACCAATGGCCAGACACATTCGCTCGCAACTTGTAGGTTCAAAGACACTTCGAGGACGGG--T	1582		
Db	295	TGACCAATGGCCAGACACATTCGCTCGCAACTTGTAGGTTCAAAGACACTTCGAGGACGGGTTT	236		
QY	1583	TTGTGGACAGAGCCCTACTTTTGGGACACGCTGAGGTGTAAGGGGACAGACAAACAGGTGA	1642		
Db	235	TTGTGGACAGAGCCCTACTTTTGGGACACGCTGAGGTGTAAGGGGACAGACAAACAGGTGA	176		
QY	1643	GGGTGAGTGTGACACCCAGAGACTGTCTTCTCGCCCTTACCTTGGCCCACTCTTACGAC	1702		
Db	175	GGGTGAGTGTGACACCCAGAGACTGTCTTCAACTTGGCCCTTACCTTGGCCCACTTACGAC	116		
QY	1703	TGGGAGCTGACAT--GACCGAGCCAGATCGTGTGACGAGGTGCTGCTC-----TGTGGC	1757		
Db	115	TGGAACCTGAAATTAGACACRCCAAAMGATCTGTGTACGAGGTCTGTCTCTGTATGACACA	56		

Oy	1758	AGCCCTGTTTATATGC-CATTACAGATGTGGTCAGACTCTTCTGGGCCCGGA	1811
Dd	55	GGACTCCTGTTTATAGCACAGATCATGTGCTCAGACTCTTCTGGGCTTAA	1
<hr/>			
RESULT	7		
LOCUS	BF038722	769 bp	mRNA linear EST 20-OCT-2000
DEFINITION	60146064IF1 NIH_MGC_66 Homo sapiens CDNA clone IMAGE:3864082 5' ,		
ACCESSION	BF038722		
VERSION	BF038722.1	GI:10745770	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.		
TITLE	NIH-MGC http://mgc.nci.nih.gov//		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgaabs@remail.nih.gov Tissue Procurement: DCRD/BMP CDNA Library Preparation: Life Technologies, Inc. DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: http://image.lnl.gov Plate: lham9605 row: a column: 11 High quality sequence stop: 632. Location/Qualifiers 1..769 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:3864082" /clone_lib="NIH_MGC_66" /tissue_type="adenocarcinoma" /lab_host="DH10B (phage-resistant)" /note="Organ: ovary; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.8 kb. Library constructed by Life Technologies."		
<hr/>			
BASE COUNT	184 a 212 c 217 g 156 t		
ORIGIN			
<hr/>			
Query Match	32.9%; Score 659.8; DB 10; Length 769;		
Best Local Similarity	97.8%; Pred. No.2.9e-113;		
Matches 701; Conservative	0; Mismatches 12; Indels 4; Gaps 3;		
Oy	930	AACCTTGAACCGCCACCAGTAGGCGAGTGAGACTGTCAGCGCGTGTGTTAGAGAGCCC	989
Dd	1	AACCTTGAACCGCCACCAGTAGGCGAGTGAGACTGTCAGCGCGTGTGTTAGAGAGCCC	60
Oy	990	AGCCCCGTGAGGTGAATCTGAGAAGCTCCGCCGCCATCCCTTCGTGATGATATGATCT	1049
Dd	61	AGCCCCGTGAGGTGAATCTGAGAAGCTCCGCCGCCATCCCTTCGTGATGATATGATCT	120
Oy	1050	CATTGCTACCTTGTGATGTGAGATCACTCCCGAAGCCCGGCGCTCCAGACTCCAGCATGGTTA	1109
Dd	121	CAATCTCTACCTTGTGATGTGAGATCACTCCCGAAGCCCGGCGCTCCAGCATGGTTA	180
Oy	1110	CTACAAAAAACCTTGTCTCTAGAGAACTCACTACATCCCAATTGAGATGTCCCAAGAAGAT	1169
Dd	181	CTACAAAAAACCTTGTCTCTAGAGAACTCACTACATCCCAATTGAGATGTCCCAAGAAGAT	240
Oy	1170	ATGCAAAAGCCCCAGAGAGAGACTCCACAGCTCTCACTGGGTGGCCAGAGCTGTGCAGACA	1229
Dd	241	ATGCAAAAGCCCCAGAGAGAGACTCCACAGCTCTCACTGGGTGGCCAGAGCTGTGCAGACA	300
Oy	1230	GCCAGATGAGAACTGTGTTGGCTTCCCTATTTTTTGTGCGGAATGCCATCTTAGGCCA	1289

DB	301	GCACATGATGAGAACTGGTGTGGCTTCCCTATTTTTCCTCCGATGTCATCTTAGGCCA	360
QY	1290	GAACACAGCCCAAAAAGGCCAGGTGAGAGTCTCTTGACAGCAAAAGTGTGTAAGACAGG	1349
DB	361	GAACACAGCCCAAAAGAGCCAGGTGAGAGTCTCTTGACAGCAAAAGTGTGTAAGACAGG	420
QY	1350	CTTGATGATGGGCTCGGTGGCCGGACAAATTCATCATCAGGCTACTGACACATGATCATCCG	1409
DB	421	CTTGATGATGGGCTCGGTGGCCGGACAAATTCATCATCAGGCTACTGACACATGATCATCCG	480
QY	1410	CCCATTTGCTGCTTGAAGCCCAAGACCAAGGTTTAAGCAGAGGGTGAGGAGACCGTGCC	1469
DB	481	CCCATTTGCTGCTTGAAGCCCAAGACCAAGGTTTAAGCAGAGGGTGAGGAGACCGTGCC	540
QY	1470	TTTCTCTTCCAGGCGCAAGCTGGACACCTTCCTGTGTGTCGTGAGAACAGTGAAGTGCAC	1529
DB	541	TTTCTCTTCCAGGCGCAAGCTGGACACCTTCCTGTGTGTCGTGAGAACAGTGAAGTGCAC	600
QY	1530	AATGGCCGAGACATATGCTCGCATCTTTAGTGTCAAGGACTGTCCAGGCGAGG--TTTG	1587
DB	601	AATGGCCGAGACATATGCTCGCATCTTTAGTGTCAAGGACTGTCCAGGCGAGGTTTGTG	659
QY	1588	GACACAGCCCTACTTTCGGGACACACCTGAGGTGTAAGGGCGACACAACAGGTGAG	1644
DB	660	GCCAGAGCCCAACTTTCGGGAGACAG-CTGAGGTGTAGGGCAGACACACAGTGCAGG	715
RESULT	8		
LOCUS	BG764118	702 bp	mRNA linear EST 15-MAY-2001
DEFINITION	602737137F1 NIH_MGC_49 Homo sapiens CDNA clone IMAGE:4862312 5',		
ACCESSION	BG764118		
VERSION	BG764118.1	GI:14074771	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	1 (bases 1 to 702)		
JOURNAL	NIH-MGC http://mgs.nci.nih.gov/ .		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)		
	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgapbs-rt@mail.nih.gov		
	Tissue Procurement: ATCC/DCTP/DRP		
	CDNA Library Preparation: Ling Hong/Rubin Laboratory		
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLB)		
	DNA Sequencing by: Incyte Genomics, Inc.		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LNLB at:		
	http://image.llnl.gov		
	Plate: LICM1722 row: j column: 09		
	High quality sequence stop: 697.		
FEATURES	Location/Qualifiers		
source	1..702		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:4862312"		
	/clone_lib="NIH_MGC_49"		
	/tissue_type="melanotic melanoma, high MDR (cell line)"		
	/lab_host="DH10B (phage-resistant)"		
	/note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2:		
	EcoRI; CDNA made by oligo-dT priming. Directionally cloned		
	into EcoRI/XhoI sites using the following 5' adaptor:		
	GGCAGCAG(G). Size-selected >500bp for average insert size		
	1.8kb. Library constructed by Ling Hong in the laboratory		
	of Gerald M. Rubin (University of California, Berkeley)		
	using ZAP-CDNA synthesis kit (Stratagene) and Superscript		
	II RT (Life Technologies). Note: this is a NIH_MGC		
	Library. !"		
BASE COUNT	183 a 180 c 202 g 136 t 1 others		

ORIGIN

Query Match 32.6%; Score 654.2; DB 10; Length 702;

Best Local Similarity 97.1%; Pred. No. 3.3e-112; Mismatches 676; Conservative 0; Indels 1; Gaps 1;

Qy	33	GGGGGGGGCTCTAGAGACCGGACCTGTACAGTTCTTTGGCTGCTGGGCCCCCTTGAG	92
Db	3	GGGGGGGGCTCTAGAGACCGGACCTGTACAGTTCTTTGGCTGCTGGGCCCCCTTGAG	62
Qy	93	TCAGCCATCATATGCTTCCGTCCTGTGTCATATCTGCTCCGACTTCTTGATCAGT	152
Db	63	TCAGCCATCATATGCTTCCGTCCTGTGTCATATCTGCTCCGACTTCTTGATCAGT	122
Qy	153	CCGGAGCTGGGGCCGCTCCACCTGCGGCGACCTTCCACTTGTGAGTGGCCATTCAGT	212
Db	123	CCGGAGCTGGGGCCGCTCCACCTGCGGCGACCTTCCACTTGTGAGTGGCCATTCAGT	182
Qy	213	CTTTGAGACAGACACAGTGGACCTGCCACAGTGGCCGAATCCAGGTTGGCAAAAGAC	272
Db	183	GTTTGAACAGACACAGTGGACCTGCCACAGTGGCCGAATCCAGGTTGGCAAAAGAC	242
Qy	273	CATTATCAATTAAGCTCTTTTGTATCTTGCCAGAGAGAGAGATGCTTGTATCAG	332
Db	243	CATTATCAATTAAGCTCTTTTGTATCTTGCCAGAGAGAGAGATGCTTGTATCAG	302
Qy	333	ATTGTTAAAGAAATGATGACATGTGACAGCCAGCTTCCAGAAAGCAAGAGAG	392
Db	303	ATTGTTAAAGAAATGATGACATGTGACAGCCAGCTTCCAGAAAGCAAGAGAG	362
Qy	393	ACGAGACAGCAGGTATCATGACACTCTGCGGGATACGCTGGAGAGCAATGCTAC	452
Db	363	ACGAGACAGCAGGTATCATGACACTCTGCGGGATACGCTGGAGAGCAATGCTAC	422
Qy	453	TGTGTATCTCTGAGAGGCTTTGGCAAGCCGAGATGCTGTCTCCACTGAAAA	512
Db	423	TGTGTATCTCTGAGAGGCTTTGGCAAGCCGAGATGCTGTCTCCACTGAAAA	482
Qy	513	GCAATATGACTTAAAGACAGCAGCAGATGAGAACCAACACACAAAGAGGGGG	572
Db	483	GCAATATGACTTAAAGACAGCAGCAGATGAGAACCAACACAAAGAGGGGG	542
Qy	573	CCGGCTAGAGAGCAAGATGAGACCATGAGACATGAGCTTCTACTCCAGACCG	632
Db	543	CCGGCTAGAGAGCAAGATGAGACCATGAGACATGAGCTTCTACTCCAGACCG	602
Qy	633	CCCTGAGCTGAGAGGATGATGATCCGAGACATGGGTGGGACAGTGGAGACAG	691
Db	603	CCCTGAGCTGAGAGGATGATGATCCGAGACATGGGTGGGACAGTGGAGACAG	662
Qy	692	TGGCTGTGCTACTGTGTCTCTCAAGAAAGTACG	727
Db	663	TGGCTGTGCTACTGTGTCTCTCAAGAAAGTACG	698

RESULT 9

BE869186 888 bp mRNA linear EST 20-OCT-2000
LOCUS 601445180F1 NIH_MGC_65 Homo sapiens cDNA clone IMAGE:3849451 5',
DEFINITION mRNA sequence.

ACCESSION BE869186
VERSION BE869186.1 GI:10317962

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

TITLE NIH-MGC http://mgc.ncl.nih.gov/

JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)

COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.

Email: cgabs-r@mail.nih.gov

Tissue Procurement: ATCC

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LNL at:

http://image.lnl.gov

Plate: LLAM9565 row: 0 column: 20

High quality sequence stop: 692.

Location/Qualifiers

1. 888

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:3849451"

/clone_1b="NIH_MGC_65"

/tissue_type="adenoecarcinoma"

/lab_host="DH10B (phage-resistant)"

/note="Organ: colon; Vector: pCMV-Sport6; Site_1: NotI;

Site_2: SalI; cloned unidirectionally. Primer: Oligo dT.

Average insert size 1.8 kb. Library constructed by Life

Technologies.

BASE COUNT 254 a 199 c 268 g 167 t

ORIGIN

Qy	309	GGAGGAGAAATGCTTGGATGAGATCTTAA- GAATGAATCGACATGTGAGACCC	367
Db	7	GGAGGAGAAATGCTTGGATGAGATCTTAA- GAATGAATCGACATGTGAGACCC	66
Qy	368	AGCTTCCCAAGAAAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	427
Db	67	AGCTTCCCAAGAAAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	126
Qy	428	ATTAGCTGGAAGAGC- AATGCTACTGTGTATCTCTGACAGAGGCTTGGCAAGCC	486
Db	127	ATTAGCTGGAAGAGC- AATGCTACTGTGTATCTCTGACAGAGGCTTGGCAAGCC	186
Qy	487	GAGATGCTGTGCTCCAGACTGAAAAAGCAGATGAATGACTTGAAGCAGCAGATG	546
Db	187	GAGATGCTGTGCTCCAGACTGAAAAAGCAGATGAATGACTTGAAGCAGCAGATG	246
Qy	547	ACCAACAGCACAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	606
Db	247	ACCAACAGCACAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	306
Qy	607	ATTAGCTTCTACTCCAGAGCAGCTCCCTGAGTGGAGAGATGATCCAGACATGG	666
Db	307	ATTAGCTTCTACTCCAGAGCAGCTCCCTGAGTGGAGAGATGATCCAGACATGG	366
Qy	667	GTGGAGAGTCCAGGCGGTGAACACAGCTGCTGTGATCTGTCTCAAGAAAGATAC	726
Db	367	GTGGAGAGTCCAGGCGGTGAACACAGCTGCTGTGATCTGTCTCAAGAAAGATAC	426
Qy	727	GAGAACTTAAAGAGAGCAGGAGGCTCAGGAGAGTGGCTGCAAGCTGAGAGAGAT	786
Db	427	GAGAACTTAAAGAGAGCAGGAGGCTCAGGAGAGTGGCTGCAAGCTGAGAGAGAT	485
Qy	787	TTGTTTCTCTCAGAGCAAGTTCAGACAGTCTACTCTGAATTGATCAGGCCAATTA	846
Db	486	TTGTTTCTCTCAGAGCAAGTTCAGACAGTCTACTCTGAATTGATCAGGCCAATTA	545
Qy	847	GAACTGAAGTCCAGGAGAGGACTTACAGAGTCTGACAAAGAAATCAGAGCTGAAA	906
Db	546	GAACTGAAGTCCAGGAGAGGACTTACAGAGTCTGACAAAGAAATCAGAGCTGAAA	605
Qy	907	AAGAAGCT- AAGATGCTGACAGAACTT- GAACCTGCCACCAAGTGGCAGTGAAGT	963
Db	606	AAGAAGCTTAAAGATGCTGAGAAACCTTTGAACCTGTCCACCAAGTGGCAGTGAAGT	665

QY 964 GTGACCCGCTGGTATTAGAGAGCCAGCCCTGT-GGAGGTGAATCTG--AGCTCCGC 1020
|||||
Db 666 GTGACCCGCTGGTATTAGAGAGCCAGCCCTGTGGGAGGTGAATCTGAAAGCTCGGC 725
QY 1021 CGGCATCTCT-TCCGTGATGATAT-TGATCTCAATGCTACCTTTGATGTGACTCCCC 1078
|||||
Db 726 CGGCATCTCTTCCGCGATGATGTGATCTCAATGCTACCTTTGATGTGACTCC 785
QY 1079 CAGCCCGGCGCTCCAGCTCCAGCATGGTTACTAGAAAACCTTGCCCTAGAGAGTCAAC 1138
|||||
Db 786 CAG-CCGGCCCTCAGCTTCCAGCATGGGTACTAGAAAACCTTTGCTGAGAGAGACATCC 844
QY 1139 A 1139
Db 845 A 845
RESULT 10
Bg120736 784 bp mRNA linear EST 30-JAN-2001
LOCUS 602347389P1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:4442038 5',
DEFINITION mRNA sequence.
ACCESSION Bg120736
VERSION Bg120736.1 GI:12614245
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 784)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM10214 row: b column: 23
High quality sequence stop: 713.
FEATURES
source
location/Qualifiers
1..784
/organism="Homo sapiens"
/db_xref="taxon:.9606"
/clone="IMAGE:4442038"
/clone_lid="NIH_MGC_90"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: Liver; Vector: PCMV-Sport6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; Oligo-dT primed.
Average insert size 1.7 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
BASE COUNT 207 a 200 c 224 g 153 t
ORIGIN

Query Match 32.4%; Score 651; DB 10; Length 784;
Best Local Similarity 93.6%; Pred. No. 1.3e-111;
Matches 735; Conservative 0; Mismatches 40; Indels 10; Gaps 5;

QY 16 ATTGAAGCAAGCGAGC-----GGGGCGCTCTACGAAGCGGACCTGTAGCAGTTTCT 70
|||||
Db 1 ATTGAAGCAAGCGAGCGGGGGCGCTCTACGAAGCGGACCTGTAGCAGTTTCT 60
QY 71 TTGGCTGCTGGGCGCTTGAGTGCAGCAGCATCATCTATCCGTCTGTGCACTATCT 130
|||||
Db 61 TTGGCTGCTGGGCGCTTGAGTGCAGCAGCATCATCTATCCGTCTGTGCACTATCT 120

QY 131 GCTCGACTTCTTGATGATCTCCCGAGAGTGGCGCCATCTGCGGCGACACCTTCC 190
|||||
Db 121 GCTCGACTTCTTGATGATCTCCCGAGAGTGGCGCCATCTGCGGCGACACCTTCC 180
QY 191 ACTTGCACTGCTTAATTCAGTCTTTTGAGAGAGCAGCAAGTGGACCTGCGCAGTCC 250
|||||
Db 181 ACTTGCACTGCTTAATTCAGTCTTTTGAGAGAGCAGCAAGTGGACCTGCGCAGTCC 240
QY 251 GAATCCAGTGGGCAAAAAGCAATTCATTAAGCTCTTCTTTGATCTTCCCGAGAGG 310
|||||
Db 241 GAATCCAGTGGGCAAAAAGCAATTCATTAAGCTCTTCTTTGATCTTCCCGAGAGG 300
QY 311 AGGAAATGCTCTTGATGATGAGATTTCTTAAGAAATGAGATGAGATGAGATGAGATG 370
|||||
Db 301 AGGAAATGCTCTTGATGATGAGATTTCTTAAGAAATGAGATGAGATGAGATGAGATG 360
QY 371 TTTCCTCAAGAAAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATG 430
|||||
Db 361 TTTCCTCAAGAAAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATG 420
QY 431 CGCTGGAAGAGAGAGATGCTGATGATGATGATGATGATGATGATGATGATGATGATG 490
|||||
Db 421 CGCTGGAAGAGAGAGATGCTGATGATGATGATGATGATGATGATGATGATGATGATG 480
QY 491 TGCTGTGCTCCACACTGAAAAGAGAGATGATGATGATGATGATGATGATGATGATGATG 550
|||||
Db 481 TGCTGTGCTCCACACTGAAAAGAGAGATGATGATGATGATGATGATGATGATGATGATG 540
QY 551 AACAGACCAAGATG 610
|||||
Db 541 AACAGACCAAGATG 599
QY 611 AGCTTCTACTCCAGATG 670
|||||
Db 600 AGCTTCTACTCCAGATG 659
QY 671 GACAGTACAGCGGTGATG 730
|||||
Db 660 GACAGTACAGCGGTGATG 717
QY 731 ATCTAAAAAGATG 790
|||||
Db 718 ATCTAAAAAGATG 775
QY 791 TTTCCT 795
Db 776 CTTCC 780

RESULT 11
Bg682548 752 bp mRNA linear EST 01-MAY-2001
LOCUS 602624362F1 NCI_CGAP_Skn4 Homo sapiens cDNA clone IMAGE:4749447 5',
DEFINITION mRNA sequence.
ACCESSION Bg682548
VERSION Bg682548.1 GI:13913945
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 752)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: James Cleaver, M.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
Clone distribution: MGC clone distribution information can be

Found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10602 row: k column: 16
High quality sequence start: 19
High quality sequence stop: 736.

FEATURES

Location/Qualifiers

1..752

source
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4749447"
/clone_lib="NCI_CGAP_Skn4"
/tissue_type="squamous cell carcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: skin; Vector: PCMV-SPORT6; Site: 1: NotI;
Site: 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.5kb. Library constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT 200 a 194 c 212 g 146 t
ORIGIN

Query Match 32.4%; Score 650.2; DB 10; Length 752;
Best Local Similarity 97.0%; Pred. No. 1.8e-111;
Matches 705; Conservative 0; Mismatches 18; Indels 4; Gaps 4;

QY 37 GGGGCTCTAGAGAGCGGACCTGTGACGTTCTTGGCTGGCCGCGGCGCTTGACATCA 96
Db 22 GGGGCTCTAGAGAGCGGACCTGTGACGTTCTTGGCTGGCCGCGG-CCCTTGAGTCA 80
QY 97 GCCATCATGCCCTATCCGCTGTGTGACATATGCTCCGACTTCTTCGATCATCTCCGC 156
Db 81 GCCATCATGCCCTATCCGCTGTGTGACATATGCTCCGACTTCTTCGATCATCTCCGC 140
QY 157 GAGTGGCCGCCATCCACTGCGGACACCTTCCACTTGCAGTGCCTAATTCAGTCTTT 216
Db 141 GAGTGGCCGCCATCCACTGCGGACACCTTCCACTTGCAGTGCCTAATTCAGTCTTT 200
QY 217 GAGACAGACCAAGTGGAGCTGGCCGACAGTGGCGAATCCAGG-TTGGCAAAAGAACAT 275
Db 201 GAGACAGACCAAGTGGAGCTGGCCGACAGTGGCGAATCCAGGTTTGGCAAAAGAACAT 260
QY 276 TATCATTAAGCTCTTCTTTGATCTTGGCCAGAGAGAGAAATGTCTTGGATCGAAT 335
Db 261 TATCATTAAGCTCTTCTTTGATCTTGGCCAGAGAGAGAAATGTCTTGGATCGAAT 320
QY 336 CTTAAAGAAATGAATGACATGTGACAGCCGCACTTCCCAAGAAACAGAGAAAGC 395
Db 321 CTTAAAGAAATGAATGACATGTGACAGCCGCACTTCCCAAGAAACAGAGAGAAAGC 380
QY 396 AGACAGCAGGTCATCATGACACTGTGGGATAGCGTGGAAAGAACGCAATGCTACTGT 455
Db 381 AGACAGCAGGTCATCATGACACTGTGGGATAGCGTGGAAAGAACGCAATGCTACTGT 440
QY 456 GGTATCTGTGACAGGCTTGGGCAAGGCCGAGATGCTGTCTCCACTTGAAAAACA 515
Db 441 GGTATCTGTGACAGGCTTGGGCAAGGCCGAGATGCTGTCTCCACTTGAAAAACA 500
QY 516 GATGAAGTACTTGAAGAGAGAGATGAGACCAAAAGCAGCAGAGAGAGGCGCG 575
Db 501 GATGAAGTACTTGAAGAGAGAGATGAGACCAAAAGCAGCAGAGAGAGGCGCGCG 560
QY 576 GCTCAGGAGCAATGATGAAGCATGTGAGCATTTGAGCTTCTACTCCAGAGCCAGCTCCC 635
Db 561 GCTCAGGAGCAATGATGAAGCATGTGAGCATTTGAGCTTCTACTCCAGAGCCAGCTCCC 620
QY 636 TGAAGTGGAGAGATGATCGAGACATGGGTGTGGACATCAGCGGTGG-AAACAGTGG 694
Db 621 TGAAGTGGAGAGATGATCGAGACATGGGTGTGGACAGCAGCGGTGGAAACAAGTGG 680
QY 695 CTGTGTACTGTGTGTCTT-CAAGAAAGATAGAGAAATTAAGAGAGCAGAGAGCC 733
Db 681 CTGTGTACTGTGTGTCTTCAAGAAAGATAGAGAAATTAAGAGAGCAGCGGAGAGGCC 740

QY 754 TCAGGGG 760
Db 741 TCCGGGG 747

RESULT 12

LOCUS BG420765 908 bp mRNA linear EST 14-MAR-2001
DEFINITION 602448558F1 NIH_MGC_14 Homo sapiens cDNA clone IMAGE:4586846 5',
mRNA sequence.
ACCESSION BG420765
VERSION BG420765.1 GI:13327271
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 908)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: DCTD/DRP
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov

plate: LLCMI316 row: d column: 15
High quality sequence stop: 719.
Location/Qualifiers
1..908
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4586846"
/clone_lib="NIH_MGC_14"
/tissue_type="renal cell adenocarcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: kidney; Vector: pOTB7; Site: 1: XhoI; Site: 2:
EcoRI; CDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5',
adaptor: GGCACGCG(G). Size-selected >800bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

FEATURES

source

BASE COUNT 235 a 229 c 274 g 170 t
ORIGIN

Query Match 31.5%; Score 632.6; DB 10; Length 908;
Best Local Similarity 89.2%; Pred. No. 3.3e-108;
Matches 770; Conservative 0; Mismatches 39; Indels 54; Gaps 6;

QY 82 GGGCCCTTGAGTCCAGCCATCATGCTATCCGTGTGTGACATATCTCCGACTTC 141
Db 1 GGGCCCTTGAGTCCAGCCATCATGCTATCCGTGTGTGACATATCTCCGACTTC 60
QY 142 TTGGATCATCTCCGCGCGGCGGCCATCCATCGCGGCGCACACCTTCCACTTGCAAGTGC 201
Db 61 TTGGATCATCTCCGCGCGGCGGCCATCCATCGCGGCGCACACCTTCCACTTGCAAGTGC 120
QY 202 CTAATTCAGTCTTTGAGACAGCAGCAAGTGGAGCTGCGCACAGTGCAGATCCAGGTT 261
Db 121 CTAATTCAGTCTTTGAGACAGCAGCAAGTGGAGCTGCGCACAGTGCAGAT----- 173
QY 262 GGCAGAAAGAACATTAATCAATTAAGCTCTTGTGATCTTCCAGAGAGAGAGAAATGTC 321
Db 174 -----CCAGAGGAGAGAAATGTC 192
QY 322 TTGGATCGAGAAATTTAAAGAAATGAACGTGACATGTGACAGCCCAAGCTTCCAGAA 381

LOCUS AL559248 645 bp mRNA linear EST 16-FEB-2001
DEFINITION AL559248 LTI_NFL008.TC2 Homo sapiens cDNA clone CSDDJ012YH21 5
prime, mRNA sequence.
ACCESSION AL559248
VERSION AL559248.1 GI:12904563
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 645)
Li, W.-B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: segrete@genoscope.cns.fr, Web: www.genoscope.cns.fr.
Location/Qualifiers
1. 645
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CSDDJ012YH21"
/clone_lib="LTI_NFL008_TC2"
/sex="male"
/tissue.type="T cells from T cell leukemia"
/note="Vector: pCMVSPORT 6; Site.1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact: Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax: (1) 301 610 8371
Email: fliang@lifestech.com URL:
http://fulllength.invitrogen.com"

BASE COUNT 167 a 174 c 177 g 127 t
ORIGIN

Query Match 31.4%; Score 629.8; DB 9; Length 645;
Best Local Similarity 98.9%; Pred. No. 1.2e-107;
Matches 634; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

42 TCACGAGCGGACCGTGTGACGATCTTGTGCTGCTGGCCCTTGTGATCCAGCAT 101
|||||
Db 5 TCACGAGCGGACCGTGTGACGATCTTGTGCTGCTGGCCCTTGTGATCCAGCAT 64
|||||

102 CATGCTATCCGCTCTGTGACATATCTGCTCGAATCTTGTGATCCTCCGCGAGT 161
|||||
Db 65 CATGCTATCCGCTCTGTGACATATCTGCTCGAATCTTGTGATCCTCCGCGAGT 124
|||||

162 GGGCCGCATCCACTGGGGCCACACCTTCCAGTGCCTAATTCAGTCTTTGAGAC 221
|||||
Db 125 GGGCCGCATCCACTGGGGCCACACCTTCCAGTGCCTAATTCAGTCTTTGAGAC 184
|||||

222 AGCACCAGTCGAGACCGCCACAGTCGGAATCCAGGTTGGCAAAAGACCATTAACA 281
|||||
Db 185 AGCACCAGTCGAGACCGCCACAGTCGGAATCCAGGTTGGCAAAAGACCATTAACA 244
|||||

282 TAAGCTCTTCTTGATCTTCCAGAGAGAGGAATGTTGGATGGAAATTTCTTAAA 341
|||||
Db 245 TAAGCTCTTCTTGATCTTCCAGAGAGAGGAATGTTGGATGGAAATTTCTTAAA 304
|||||

342 GAATGAATGAGACATGTGAGAGCCAGCTTCCAGAAAGACAGAGAGAGAGAG 401
|||||
Db 305 GAATGAATGAGACATGTGAGAGCCAGCTTCCAGAAAGACAGAGAGAGAGAGAG 364
|||||

402 CCAGGTATCATGACACTGTGGGGATAGCTGGAAAGAGCAATCTCTGCGTATC 461
|||||
Db 365 CCAGGTATCATGACACTGTGGGGATAGCTGGAAAGAGCAATCTCTGCGTATC 424
|||||

462 TCTGACAGCAGCCCTGGGGCAAGCCGAGATGCTGTCTCCACACTGAAAAAGAGATGAA 521
|||||

|||||
Db 425 TCTGCACGACGCGCTGGGGCAAGGCCGAGATGCTGTGCTCCACACTGAAAAAGAGATGAA 484
|||||

522 GTACTTAGACGACGACGAGATGAGACCAACAGACACAGAGAGGCGCGGCTCAG 581
|||||
Db 485 GTACTTAGACGACGACGAGATGAGACCAACAGACACAGAGAGGCGCGGCTCAG 544
|||||

582 GAGCAAGATGAGACGACGAGATGAGCTTCTCTCCAGAGCCAGCTCCCTGAGGT 641
|||||
Db 545 GAGCAAGATGAGACGACGAGATGAGCTTCTCTCCAGAGCCAGCTCCCTGAGGT 604
|||||

642 GGAAGAGATGATCCGAGACATGGGTGTGGACAGTCACGCG 682
|||||
Db 605 GGAAGAGATGATCCGAGACATGGGTGTGGACAGTCACGCG 645
|||||

RESULT 15
BE546959
LOCUS BE546959 825 bp mRNA linear EST 09-AUG-2000
DEFINITION 601071929F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:3458023 5',
mRNA sequence.
ACCESSION BE546959
VERSION BE546959.1 GI:9775604
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 825)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapds-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: Incyte Genomics, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
plate: LHAM8448 row: j column: 08
High quality sequence stop: 655.
Location/Qualifiers
1. 825
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3458023"
/clone_lib="NIH_MGC_12"
/tissue.type="cervical carcinoma cell line"
/lab_host="DH10B"
/note="Organ: cervix; Vector: pCMV-SPORT6; Site.1: NotI;
Site.2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.4 kb. Library prepared by Life
Technologies."

BASE COUNT 230 a 209 c 242 g 144 t
ORIGIN

Query Match 31.3%; Score 628; DB 10; Length 825;
Best Local Similarity 95.6%; Pred. No. 2.4e-107;
Matches 711; Conservative 0; Mismatches 25; Indels 8; Gaps 6;

126 TATCTGCTCCGACTTCTTGATCAGTCCGCGAGTGGCCGATCCAGTGGCGCACAC 185
|||||
Db 1 TATCTGCTCCGACTTCTTGATCAGTCCGCGAGTGGCCGATCCAGTGGCGCACAC 60
|||||

186 CTTCACCTTCAGGCGCTTAATTCAGTCTTGGACAGCAAGTGGAGCTGGCCACA 245
|||||
Db 61 CTTCACCTTCAGGCGCTTAATTCAGTCTTGGACAGCAAGTGGAGCTGGCCACA 120
|||||

246 GTGCGAATCCAGTTGGCAAAAGACATTAACAATGCTCTTGTATCTTGCCCA 305
|||||

Db	121	GTGCCAAATTCACAGGTGTGGCAAAAAGAACCAATATTCANTAAAGCTCTTCTTGATCTTGGCCA	180
Qy	306	GGAGGAGGAGGAATGCTTTGGATTCGAGAAATTTCTTAAAGATGAACCTGGACAAATGTCAAGCC	365
Db	181	GGAGGAGGAGGAATGCTTTGGATTCGAGAAATTTCTTAAAGATGAACCTGGACAAATGTCAAGCC	240
Qy	366	CCAGCTTTCCCAAGAAAGACAGAGAGAGAAACAGACAGCCAGCCAGGCATCATGCAGACCTGGG	425
Db	241	CCAGCTTTCCCAAGAAAGACAGAGAGAGAAACAGACAGCCAGGCATCATGCAGACCTGGG	300
Qy	426	GGATACGCTTGGAAAGCCAAATGCTACTGTGTATCTCTGACAGCAGAGCCTTGGGCAAGCC	485
Db	301	GGATACGCTTGGAAAGCCAAATGCTACTGTGTATCTCTGACAGCAGAGCCTTGGGCAAGCC	360
Qy	486	CGAGATGCTGTGCTCCACACTGTAAGAAAGACAGATGAAGTACTTAAGACAGACAGCAGATGA	545
Db	361	CGAGATGCTGTGCTCCACACTGTAAGAAAGACAGATGAAGTACTTAAGACAGACAGCAGATGA	420
Qy	546	GACCAAAACAGACACAAGAGAGAGGCGGGCGGCTCAGAGACCAAGATGAAGCAGATGAGAGA	605
Db	421	GACCAAAACAGACACAAGAGAGAGGCGGGCGGCTCAGAGACCAAGATGAAGCAGATGAGAGA	480
Qy	606	GATTGAGCTTCTACTCTCCAGAGCAGCTCCCTGAGGTGGAGAGATGATCCGAGACATGGG	665
Db	481	GA-TGAGCTTCTACTCTCCAGAGCAGCTCCCTGAGGTGGAGAGATGATCCGAGACATGGG	539
Qy	666	TGTGGACACGTACAGCGGT-GGAACAGCTGGCTGTGATCTGTCTCTCAAGAAAGAT	724
Db	540	TGTGGACACGTACAGCGGTGGGAACAGCTGTGTGTACTGTGTCTCTCAAGAAAGAT	599
Qy	725	ACGAGATCTTAAAGAGAGCAGGAGAGGCTT-CAGGGAGAGTGGCTGACAACTGAGGAAG	783
Db	600	ACGAGATCTTAAAGAGAGCAGGAGAGGCTTCCAGGGAGAGTGGCTGACAACTGAGGAAG	659
Qy	784	GATTGTCTTCTCCAGAAAGCAAGTTGACAGACAGTCTACTCTGAATTTGATCAGGCCAAG	843
Db	660	GA-TGCTTCTCTCCAGAAAGCAAGTTGACAGACAGTCTACTCTGAATTTGATCAGGCCAAG	714
Qy	844	TTAGAACTGAAGTCAAGCCAGAGAG	867
Db	715	GTAGAACTGAAGTCAAGCCAGAGAG	738
RESULT	16		
BE889971			
LOCUS		983 bp	mRNA
DEFINITION	601512330P1 NIH_MGC_71 Homo sapiens	cdna clone IMAGE:3913854 5',	
ACCESSION	BE889971		
VERSION	BE889971.1	GI:10347827	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	1 (bases 1 to 983)		
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)		
	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: c9apbs-rfemail.nih.gov		
	Tissue Procurement: ATCC		
	CDNA Library Preparation: Life Technologies, Inc.		
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)		
	DNA Sequencing by: Incyte Genomics, Inc.		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LNL at:		
	http://image.llnl.gov		
	Plate: LLM9734 row: k column: 07		
	High quality sequence stop: 664.		
	Location/Qualifiers		
FEATURES			
SOURCE	1. .983		
	/organism="Homo sapiens"		

DB	Query Match	Best Local Similarity	Score	DB	Length
31.2%	92.9%	626.8	10	983	
Matches	691	Conservative	0	Mismatches	47
				Indels	6
				Gaps	3
BASE COUNT	279 a	259 c	289 g	156 t	
ORIGIN					
Query Match	31.2%	Score 626.8	DB 10	Length 983	
Best Local Similarity	92.9%	Pred No. 46-107			
Matches	691	Conservative	0	Mismatches	47
				Indels	6
				Gaps	3
BASE COUNT	279 a	259 c	289 g	156 t	
ORIGIN					
Query Match	31.2%	Score 626.8	DB 10	Length 983	
Best Local Similarity	92.9%	Pred No. 46-107			
Matches	691	Conservative	0	Mismatches	47
				Indels	6
				Gaps	3
BASE COUNT	279 a	259 c	289 g	156 t	
ORIGIN					
Query Match	31.2%	Score 626.8	DB 10	Length 983	
Best Local Similarity	92.9%	Pred No. 46-107			
Matches	691	Conservative	0	Mismatches	47
				Indels	6
				Gaps	3
BASE COUNT	279 a	259 c	289 g	156 t	
ORIGIN					
Query Match	31.2%	Score 626.8	DB 10	Length 983	
Best Local Similarity	92.9%	Pred No. 46-107			
Matches	691	Conservative	0	Mismatches	47
				Indels	6
				Gaps	3
BASE COUNT	279 a	259 c	289 g	156 t	
ORIGIN					
Query Match	31.2%	Score 626.8	DB 10	Length 983	
Best Local Similarity	92.9%	Pred No. 46-107			
Matches	691	Conservative	0	Mismatches	47
				Indels	6
				Gaps	3
BASE COUNT	279 a	259 c	289 g	156 t	
ORIGIN					
Query Match	31.2%	Score 626.8	DB 10	Length 983	
Best Local Similarity	92.9%	Pred No. 46-107			
Matches	691	Conservative	0	Mismatches	47
				Indels	6
				Gaps	3
BASE COUNT	279 a	259 c	289 g	156 t	
ORIGIN					
Query Match	31.2%	Score 626.8	DB 10	Length 983	
Best Local Similarity	92.9%	Pred No. 46-107			
Matches	691	Conservative	0	Mismatches	47
				Indels	6
				Gaps	3
BASE COUNT	279 a	259 c	289 g	156 t	
ORIGIN					
Query Match	31.2%	Score 626.8	DB 10	Length 983	
Best Local Similarity	92.9%	Pred No. 46-107			
Matches	691	Conservative	0	Mismatches	47
				Indels	6
				Gaps	3
BASE COUNT	279 a	259 c	289 g	156 t	
ORIGIN					
Query Match	31.2%	Score 626.8	DB 10	Length 983	
Best Local Similarity	92.9%	Pred No. 46-107			
Matches	691	Conservative	0	Mismatches	47
				Indels	6
				Gaps	3
BASE COUNT	279 a	259 c	289 g	156 t	
ORIGIN					
Query Match	31.2%	Score 626.8	DB 10	Length 983	
Best Local Similarity	92.9%	Pred No. 46-107			
Matches	691	Conservative	0	Mismatches	47
				Indels	6
				Gaps	3
BASE COUNT	279 a	259 c	289 g	156 t	
ORIGIN					
Query Match	31.2%	Score 626.8	DB 10	Length 983	
Best Local Similarity	92.9%	Pred No. 46-107			
Matches	691	Conservative			

mRNA sequence.
 ACCESSION BE219707
 VERSION BE219707.1 GI:11125801
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 632)
 NIH-MGC http://mgc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: c9apbs-remail.nih.gov
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov
 Plate: IRAL row: e column: 07
 High quality sequence stop: 630.
 Location/Qualifiers
 1..632
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2821007"
 /clone_1lb="NIH-MGC-7"
 /tissue_type="small cell carcinoma"
 /cell_line="MGC3"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: lung; Vector: POT87; Site_1: XhoI; Site_2: EcoRI; CDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAC(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."
 BASE COUNT 163 a 171 c 172 g 126 t
 ORIGIN
 Query Match 30.3%; Score 608.6; DB 10; Length 632;
 Best Local Similarity 98.6%; Pred. No. 1.1e-103;
 Matches 614; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
 Oy 42 TCTACGAAGCGGACCTGTAGCACTTTCTTGCTGCTGGCCCTTGAGTCCAGCCAT 101
 |||||||
 Db 7 TCTACGAAGCGGACCTGTAGCACTTTCTTGCTGCTGGCCCTTGAGTCCAGCCAT 66
 Oy 102 CATGCCATCGGCTGCTGTGCACTATCTGCTCGACTTCTTGATCATGCCGCGAGCT 161
 |||||||
 Db 67 CATGCCATCGGCTGCTGTGCACTATCTGCTCGACTTCTTGATCATGCCGCGAGCT 126
 Oy 162 GGGCGCCATCCACTGCGGCGACACACTTCCACTTGCTTAATTCAGTCCCTTGAGAC 221
 |||||||
 Db 127 GGGCGCCATCCACTGCGGCGACACACTTCCACTTGCTTAATTCAGTCCCTTGAGAC 186
 Oy 222 AGACCAAGTGGACCTGCCACAGTCCGATCCAGGTTGGCAAAAGAACCATTAATCAA 281
 |||||||
 Db 187 AGACCAAGTGGACCTGCCACAGTCCGATCCAGGTTGGCAAAAGAACCATTAATCAA 246
 Oy 282 TAAGCTCTTCTTGATCTTGCCGAGGAGGAGGAGATGCTTGATGAGATCTTAATA 341
 |||||||
 Db 247 TAAGCTCTTCTTGATCTTGCCGAGGAGGAGGAGATGCTTGATGAGATCTTAATA 306
 Oy 342 GAATGAACGACAAATGTCAAGGCCAGCTTTCCAGAAAAGCAAGAGAGAAAGAGACAG 401
 |||||||
 Db 307 GAATGAACGACAAATGTCAAGGCCAGCTTTCCAGAAAAGCAAGAGAGAAAGAGACAG 366
 Oy 402 CCAGGTATCATCGACACTCTGGGGATPAGCTGGAAGAACGCAATCTACTGTGTATC 461
 |||||||
 Db 367 CCAGGTATCATCGACACTCTGGGGATPAGCTGGAAGAACGCAATCTACTGTGTATC 426

Oy 462 TCTGCAGCAGACCTTGGGCAAGCCGAGATGCTGTGCTCCACTGAAGAAAGAGATGAA 521
 |||||||
 Db 427 TCTGCAGCAGACCTTGGGCAAGCCGAGATGCTGTGCTCCACTGAAGAAAGAGATGAA 486
 Oy 522 GTACTTAGACGACGAGGATGAGACCAACAGCAAGAGAGGCGGCGGCTCAG 581
 |||||||
 Db 487 GTACTTAGACGAGAGGATGAGACCAACAGCAAGAGAGGCGGCGGCTCAG 546
 Oy 582 GAGCAGATGAAGACCATGAGAGGATTTGAGCTTCTACTCCAGAGCCAGCTCCAGGT 641
 |||||||
 Db 547 GAGCAGATGAAGACCATGAGAGGATTTGAGCTTCTACTCCAGAGCCAGCTCCAGGT 606
 Oy 642 GGAGAGATGATCCGAGACATGG 664
 |||||||
 Db 607 GGAGAGATGATCCGAGACTGG 629
 RESULT 18
 BG117979 912 bp mRNA linear EST 30-JAN-2001
 LOCUS BG117979
 DEFINITION 602351223F1 NIH-MGC_90 Homo sapiens cDNA clone IMAGE:4446149 5',
 mRNA sequence.
 ACCESSION BG117979
 VERSION BG117979.1 GI:12611485
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 912)
 NIH-MGC http://mgc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: c9apbs-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov
 Plate: LLAM10224 row: n column: 06
 High quality sequence stop: 678.
 Location/Qualifiers
 1..912
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4446149"
 /clone_1lb="NIH-MGC-90"
 /tissue_type="adenocarcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: liver; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; cloned unidirectionally; oligo-dT primed. Average insert size 1.7 kb. Library enriched for full-length clones and constructed by Life Technologies.
 Note: this is a NIH-MGC library."
 BASE COUNT 236 a 243 c 273 g 160 t
 ORIGIN
 Query Match 29.9%; Score 599.4; DB 10; Length 912;
 Best Local Similarity 91.3%; Pred. No. 5.2e-102;
 Matches 737; Conservative 0; Mismatches 56; Indels 14; Gaps 9;
 Oy 14 AAATTGAAGCAAGCGAGCGGGCGC-----CTGTAGCAAGCCGAGCTGTAGCAG-TTT 68
 |||||||
 Db 2 AAATTGAAGCAAGCGAGCGGGCGCGGTGTAGCAAGCCGAGCTGTAGCAGATT 61
 Oy 69 CTTTGAGCTGCGGCGCCCTTGAGTCCAGCATCATGCTATCCGTCTGTGCACTAT 128
 |||||||
 Db 62 CTTTGAGCTGCGGCGCCCTTGAGTCCAGCATCATGCTATCCGTCTGTGCACTAT 121

OY	129	CTGCGCCAGCTCTTGAGTCACTCCCGGACGTGGCGCCGATCCACTGCGGCACACCTT	188
OY	129		
Db	122	CTGCTCCGACTCTTCTTGATCACTCCCGGACGTGGCGCATCCACTGCGGCACACCTT	181
OY	189	CCACTTGCAGTCCCTAATTCACTCCCTTGTGAGACAGCACACCAGTCGACCTGGCCACAGTG	248
Db	182	CCACTTGCAGTCCCTAATTCACTGGTGTGTGAGACAGCACCAACTCGAAGCTGGCCACAGTG	241
OY	249	CCGAATCCAGGTTGGCAAAAGAACCATTTATCAATTAAGCTCTTCTTTGATCTTGGCCAGGA	308
Db	242	CCGAATCCAGGTTGGCAAAAGAACCATTTATCAATTAAGCTCTTCTTGAATCTTGGCCAGGA	301
OY	309	GGAGAGAAATGCTTTGAGTCAAGTATTCCTTAAAGAAATGAATGCAATGTCAGAGCCCA	368
Db	302	GGAGAGAAATGCTTTGAGTCAAGTATTCCTTAAAGAAATGAATGCAATGTCAGAGCCCA	361
OY	369	-GCTTTCACAGAACAGACAGAGAACGACAGCCAGGTCAATCGACACTTGGCGG	427
Db	362	GCTTTCACAGAACAGAGAACGAGAACGACAGGTCAATCGACACTTGGCGG	421
OY	428	ATACGCTGGAACAGCAATGCTACTGTTGTTCTCTGTGACGAGGACTTTGGGCAAGG-CC	486
Db	422	ATACGCTGGAACAGCAATGCTACTGTTGTTCTCTGTGACGAGGACTTTGGGCAAGGCC	481
OY	487	GAGATGCTGTGCTCCACACTGAAAAAGCAGATGAAGTACTTTAGACAGCAGCAGGATGAG	546
Db	482	GAGATGCTGTGCTCCACACTGAAAAAGCAGATGAAGTACTTTAGACAGCAGCAGGATGAG	541
OY	547	ACCAACAAAGCACAAGAGAGGCGGCGGCTCAGAGACGAAGATGAAGCACTTGAAGCAG	606
Db	542	ACCAACAAAGCACAAGAGAGGCGGCGGCTCAGAGACGAAGATGAAGCACTTGAAGCAG	601
OY	607	ATTGAGCTTCTACT-CCAGAGCCAGCTCCCTGAGGTGAGAGAG-ATGATCCAGACAT-G	663
Db	602	ATTGAGCTTCTACTCCACAGAGCCAGCGCCGTGAGGTGGGGGAGCATGATCCGACACATGG	661
OY	664	GGTGTGGGACATCAGCGGT--GGAACAGCTGGCTGTGTACTGTGTCTCTCAAGAAAG	721
Db	662	GGTGTGGGACATCAGCGGTGGGAACCAAGCTGGTGTGTACTGCGTGGTCTCTCCAGAAAG	721
OY	722	AGTACGAAATCTTAAAGAGGACGCAAGAGCCCTCAGGGGATGGCTGCAAGCTGAGAA	781
Db	722	AGCAGGACATCTACAGAGGCCCGGCGAGGCTCTCAGAGAAAGTGGGTGCACACCCGGAG	781
OY	782	AGGATTTGTTTTCTTCACAGAACAGT	808
Db	782	AGGATTT-CTTTCACAGAACAGT	806

RESULT	19
AL529568	
LOCUS	
DEFINITION	AL529568 LTI_NFL001_NBCA Homo sapiens cDNA clone CSDD006YH09 5 prime, mRNA sequence.
ACCESSION	AL529568
VERSION	AL529568.1 GI:12793061
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 932)
AUTHORS	Ll.N.B., Gruber,C., Jesssee,J. and Polayes,D.
TITLE	Full-length cdna libraries and normalization Unpublished (2001)
JOURNAL	Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - France Email: seqret@genoscope.cns.fr, Web : www.genoscope.cns.fr.
COMMENT	Location/Qualifiers 1..932
FEATURES	/organism="Homo sapiens"
SOURCE	

```

BASE COUNT      226 a      253 c      242 g      200 t      11 others
ORIGIN
/dB_xref="taxon:3606"
/clone="CS0DD006YH09"
/clone_lib="LRI_NFL001_NB04"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
/note="Organ: brain; Vector: PCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dt) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the PCMVSPORT 6
vector. Library was normalized. Library was constructed
by Life Technologies. Contact: Feng Liang Life
Technologies, a division of Invitrogen 9800 Medical Center
Drive Rockville, Maryland 20850, USA Fax: (1) 301 610
8371 Email: liang@lifetech.com URL:
http://fulllength.invitrogen.com

```

[illegible]

```

ACCESSION AL580903 prime, mRNA sequence.
VERSION AL580903.1 GI:12947374
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 619)
JOURNAL Full-length cDNA libraries and normalization
COMMENT Unpublished (2001)
CONTACT: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
source
1..619
location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_1ib="CS0DJ012YH21"
/clone_1ib="LTI_NFL008_Tc2"
/sex="male"
/lissue_type="T cells from T cell leukemia"
/note="Vector: PCWVSPORT 6; Site_1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the PCWVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

BASE COUNT 134 a 171 c 176 g 136 t 2 others
ORIGIN
Query Match 27.7%; Score 556; DB 9; Length 619;
Best Local Similarity 97.7%; Pred. No. 6.8e-94;
Matches 604; Conservative 2; Mismatches 7; Indels 5; Gaps 4;

QY 1285 GGGCAGAAACAGCCCAAAAGCCCGAGTCCCTCTTGACGAAAGATGTGTAAAG 1344
DB 619 GGGCAGAAACAGCCCAAAAGCCCGAGTCCCTCTTGACGAAAGATGTGTAAAG 560
QY 1345 ACAAGCTTCAGTGGGCTGGTGGCCGCAAAATTCATCCAGCCTACTGACACAGTCATG 1404
DB 559 ACAAGCTTCAGTGGGCTGGTGGCCGCAAAATTCATCCAGCCTACTGACACAGTCATG 500
QY 1405 ATCCGCCCATTTGCTGTAAAGCCCAAGGTTAAGCAGAGGGTGAAGGAGTGAAGACC 1464
DB 499 ATCCGCCCATTTGCTGTAAAGCCCAAGGTTAAGCAGAGGGTGAAGGAGTGAAGACA 440
QY 1465 GTGCTCTCTCTTCCAGGCAAGCTGACACCTTCCTGTGTCGTGAGAAAGTGAAGTC 1524
DB 439 GTGCTCTCTCTTCCAGGCAAGCTGACACCTTCCTGTGTCGTGAGAAAGTGAAGTC 380
QY 1525 TGACCAATGGCCAGACACATGCTGCAACTTGTAGTCAAGACTGTCCAGGACAGG -T 1582
DB 379 TGACCAATGGCCAGACACATGCTGCAACTTGTAGTCAAGACTGTCCAGGACAGG 320
QY 1583 TTGTGACACAGAGCCCTACTTTGGGAGCAGCTGAGGTGTAAAGGACAGACAAAGGTGA 1642
DB 319 TTGTGACACAGAGCCCTACTTTGGGAGCAGCTGAGGTGTAAAGGACAGACAAAGGTGA 260
QY 1643 GGGTGAAGTGAACACCCAGAGAGAGTCTTCGCTCCCTCCACCCGACCCACTCTATAGAC 1702
DB 259 GGGTGAAGTGAACACCCAGAGAGAGTCTTCGCTCCCTCCACCCGACCCACTCTATAGAC 200
QY 1703 TGGAGAGTGAACATGACAGCCCACTGATCTGTACAGAGAGTCTCTGCT -CTGTTGCCAGGC 1761
DB 199 TGGAGAGTGAACATGACAGCCCACTGATCTGTACAGAGAGTCTCTGCTGTTGCCAGGC 140

```

```

QY 1762 TCTTGTATTATAGCATGATCATGATGTCAGACTCTTCTGGGCTCG -AGACCAAGGT 1820
DB 139 TCTGTATTATAGCATGATCATGATGTCAGACTCTTCTGGGCTCGGAGAAACCAAGGT 80
QY 1821 CACTTGTGACTGTCTCTGTGACACAGAGTCTTGAGGACATCTCAGGACGCTCAGCCCA 1880
DB 79 CACTTGTGACTGTCTCTGTGACACAGAGTCTTGAGGACATCTCAGGACGCTCAGCCCA 20
QY 1881 AGC-TTCTACTGTGCTTT 1897
DB 19 AGCGTTCTACTGTGCTTT 2

RESULT 21
AL582265/c 595 bp mRNA linear EST 16-FEB-2001
LOCUS AL582265 LTI_NFL010_BC2 Homo sapiens cDNA clone CS0DL005YM09 3
DEFINITION prime, mRNA sequence.
ACCESSION AL582265
VERSION AL582265.1 GI:12950078
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 595)
JOURNAL Full-length cDNA libraries and normalization
COMMENT Unpublished (2001)
CONTACT: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
source
1..595
location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DL005YM09"
/clone_1ib="LTI_NFL010_BC2"
/sex="male"
/lissue_type="B cells from Burkitt lymphoma"
/note="Vector: PCWVSPORT 6; Site_1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the PCWVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

BASE COUNT 113 a 162 c 165 g 150 t 5 others
ORIGIN
Query Match 25.9%; Score 520.4; DB 9; Length 595;
Best Local Similarity 97.0%; Pred. No. 3e-87;
Matches 578; Conservative 5; Mismatches 7; Indels 6; Gaps 5;

QY 1076 CCCGAGCCGCGCCCTCAGCTCCAGACATGTTACTATAC -GAAAACTTTGCTTGAAGAG 1134
DB 595 CCCGAGCCGCGCCCTCAGCTCCAGACATGTTACTATACGGAAGAACTTTGCTTGAAGAG 536
QY 1135 TCACACTCCCC -AATTGAGATGTCC -AAGAAGATATGCAAAAGGCCCCAGGAGAGAT 1192
DB 535 TCACACTCCCAATATCAGATGTCCCAAGAAAGATATGCAAAAGGCCCCAGGAGAGAT 476
QY 1193 CCCAGCTCTACTAGTGGTGGCCAGAGCTGTGCAGAGAGCCAGATGAGAGAACTGTTGGTG 1252
DB 475 CCCAGCTCTACTAGTGGTGGCCAGAGCTGTGCAGAGAGCCAGATGAGAGAACTGTTGGTG 417
QY 1253 CATTCCCTATTTTTGTCCGCAATGCATCTTAGGCCAGAAACAGCCCAAAAGGCCAGGT 1312

```

```
|||||
Db 416 CTTCCCTATTTTTGTCCGGAATGCATCCTAGGCCAGAAACAGCCCAAGAGGCCAGGT 357
OY 1313 CAGAGTCCTTTCAGCAAGATGCTGTAGAGCAGAGCTTCATGGCTCGTGGCCGA 1372
Db 356 CAGAGTCCTTTCAGCAAGATGCTGTAGAGCAGAGCTTCATGGCTCGTGGCCGA 297
OY 1373 CAAATTCATCCAGCTTACTGACACAGTCATGATCCGCCATTCGCTTTAACCCAGA 1432
Db 296 CAAATTCATCCAGCTTACTGACACAGTCATGATCCGCCATTCGCTTTAACCCAGA 237
OY 1433 CCAAGTTAAAGCAGAGGTTAGGTTGAAGACCTGCTTCCTTCAGGCCAAGCTGG 1492
Db 236 CCAAGTTAAAGCAGAGGTTAGGTTGAAGACCTGCTTCCTTCAGGCCAAGCTGG 177
OY 1493 ACACCTTCCTGCTGCTGAGACAGTGTGACCAATGGCCAGACATGCTGCA 1552
Db 176 ACACCTTCCTGCTGCTGAGAACAGTGTGACCAATGGCCAGACATGCTGCA 117
OY 1553 CTTGTAGTCAGAGCTGTCCAGCAGAGG--TTTGTGACAGAGCCCTACTTCGGGACC 1610
Db 116 CTTGTAGTCAGAGCTGTCCAGCAGAGGTTTGTGACAGAGCCCACTTCGGGACC 57
OY 1611 AGCCTGAGTGTAAAGGCGACACAAGAGTGAAGGTTGATGTGACACCCAGAGACT 1666
Db 56 AGCCTGAGTGTAAAGGCGAACAACAGAGTGAAGTGTGACACCCAGAGACT 1

RESULT 22
BI151643 702 bp mRNA linear EST 05-JUL-2001
LOCUS 602916189F1 NCI_CGAP_Lu29 Mus musculus cDNA clone IMAGE:506423 5',
DEFINITION mRNA sequence.
ACCESSION BI151643
VERSION BI151643.1 GI:14611644
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 702)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgrabbs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM11179 row: b column: 24
High quality sequence stop: 694.
Location/Qualifiers
1. 702
/organism="Mus musculus"
/strain="C2BEH II (fetal)"
/db_xref="taxon:10090"
/clone="IMAGE:506423"
/clone_lib="NCI_CGAP_Lu29"
/tissue_type="spontaneous tumor, metastatic to mammary.
stem cell origin"
/lab_host="DH10B"
/note="Organ: lung; Vector: PCMV-SPORE; Site_1: Salt;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"

BASE COUNT 186 a 182 c 190 g 144 t

ORIGIN
```

Query Match 24.8%; Score 498; DB 10; Length 702;
Best Local Similarity 87.4%; Pred. No. 4.3e-83;
Matches 568; Conservative 0; Mismatches 80; Indels 2; Gaps 2;

```
OY 72 TGGTGCTGCGGCCCTTGAAGTCCAGCATATGCTATCCGTCGTCTGTGACATATC 131
Db 41 TGGTGCTGCGGCCCTTGAAGTCCAGCATATGCTATCCGTCGTCTGTGACATATC 100
OY 132 CTCGCACTTCTTGATACATCCCGGACGAGCGGCATGCTATGCGGCACACTTCCA 191
Db 101 CTCGCACTTCTTGATACATCCCGGACGAGCGGCATGCTATGCGGCACACTTCCA 160
OY 192 CTTGCACTGCTTAATTCAGTCTTTGAGACAGCAGCAAGTGGACCTGCCACAGTCCG 251
Db 161 TCTGCAATGCTTAATTCAGTGTGAGACAGCAGCAAGTGGACCTGCCACAGTGTAG 220
OY 252 AATTCAGGTTGGCAAAAGAACATTAATCAATAGCTCTTGTGATCTTCCAGAGGA 311
Db 221 AATTCAGGTTGGCAAAAGAACATTAATCAATAGCTCTTGTGATCTTCCAGAGGA 280
OY 312 GGAGAAATGCTTGGATGAGCAATTTCTTAAAGAAATGAATGAACTGACAGCCAGCT 371
Db 281 GGAGAAATGCTTGGATGAGCAATTTCTTAAAGAAATGAATGAACTGACAGCTCA -AGCTCAGCT 339
OY 372 TTCCAGAAAGACAGAGAGAAACGAGACAGCGATCATCATGACACTCTCGGGATAC 431
Db 340 TTCCAGAAAGACAGAGAGAAACGAGACAGCGCATATGACACTCTCAGCGGACAC 399
OY 432 GCTGGAAGAACCAATGCTACTGTGATATCTCTGACAGAGCCTTGGGCAAGCCAGAT 431
Db 400 CCTGGAAGAACCAATGCTACTGTGATATCTCTGACAGAGCCTTAAACAAAGCAGAGAT 459
OY 492 GCTGTGCTCCACACATGAAAAAGAGATGAATGAACTTAAGACAGCAGAGATGAGACAA 551
Db 460 GCTGTGCTCCACCTGAAAAAACAGATGAATGAACTTAAGACAGCAGAGATGAGACAA 519
OY 552 ACAAGCACAAGAGAGAGCGCGGCTCAGAGCAAGATGAAGACATGAGCAGAGATGA 611
Db 520 ACAAGCACAAGAGAGAGCGCGGCTCAGAGCAAGATGAAGACATGAGCAGAGATGA 579
OY 612 GCTTCTACTCCAGAGCAGCTCCCTGAGTGAGAGATGATCCGAGCATGGGTGGG 671
Db 580 GCTTCTACTCCAGAGCAGCGGTCTGAGTGAGAGATGATCCGAGCATGGGTGGG 639
OY 672 AC-AGTCAGCGGTGGAACACAGCTGCTGTGACGTGTCTCAAGAAA 720
Db 640 ACAAGTCAGCGGTGGAACACAGCTGCTGTGACGTGTCTCAAGAAA 689

RESULT 23
AL563489 905 bp mRNA linear EST 16-FEB-2001
LOCUS AL563489/C
DEFINITION prime, mRNA sequence.
ACCESSION AL563489
VERSION AL563489.1 GI:12912932
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 905)
Li, W.B., Gruber, C., Jessup, J., and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
Location/Qualifiers
1. 905
/organism="Homo sapiens"
/db_xref="taxon:9606"

FEATURES
source
```

```

/clone="CS0DD006YH09"
/clone_lib="LTI_NFL001_NBC4"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
/note="Organ: brain; Vector: PCWVSORT 6; 1st strand cDNA
was primed with a Noli-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the PCWVSORT 6
vector. Library was normalized. Library was constructed
by Life Technologies, a division of Feng Liang Life
Technologies, a division of Invitrogen 9800 Medical Center
Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
8371 Email : liang@lifetech.com URL :
http://fulllength.invitrogen.com"

```

Query Match	24.4%	Score 490.4	DB 9	Length 905
Best Local Similarity	91.1%	Pred. No. 1.1e-81		
Matches 560	Conservative 4	Mismatches 46	Indels 5	Gaps 4

Qy	1382	TCCAGCCTACTGCACACAGTCATGATCCGCCCATTTGCTGTTAAAGCCCAAGACCAAGCTTA	1441
Db	616	TTCTTCCAACTAAAAAGAGTCATGACGCCGCCCATTTGCTGTTAAAGCCCAAGACCAAGCTTA <td>557</td>	557

QY 1442 AGCAGAGGGTGAGGGTGACAGACCGCTGCTTCTCTCT-TCACGGCCAAAGCTGGACACCTTC 1500
 Db 556 AGCAGAGGGTGAGGGTGAACAGTGCCTCTCTCTCTCTCCAGGCACAGCTGGACACCTTC 497

QY	1501
CTGTGGTCTGTGACACAGTGTGACTGTGTGACCAATGGCCAGACACATGCTTGCACACTTGTAGG	1560
DB	496
CTGTGGTCTGTGACACAGTGTGACTGTGTGACCAATGGCCAGACACATGCTTGCACACTTGTGTGG	437

OY	1561	TC AAGGACTGTCCAGGACGGG--TTTGTGACAGAGCCCTACTTTTCGGGACACAGCCTGTAG	1618
Db	436	TC CAGGACTGTCCAGGACGGGTTTGTGCACAGAGCCCACTTTCGGGACCCAGCCTGTAG	377

QY 1739 CAGGTCCGCT-CTGTGTCAGAGCTCTGTATTATAGCATGATCAGATGTGTCAGACTC 1797

Db 256 CAGGTCCGCTCTGTGTCACAGCTCTCTGTTTATAGCATGATGATGTGTCAAATCTC 197

QY 1798 TTTCGTGGGCGCTGGAGACCAAGGTCACTGTGTGACTGTCTCTGTGGACACAGAGTGTGAG 1857

Db 196 TTTCTGGGCGCTGGAGACCAAGGTCACTGTGTGACTGTCTCTGTGGACACAGAGTGTGAG 137

QY 1858 GCATCTCAGGCGCCTCAGCCCAAGCTCTACCTGCCTTGACTTGCTCTAGCATAGCC 1917
 136 GCTTCTCAGGCGACCCCTGAGCCCGAGGCTCTCTCCCTGACCTTTCATTTGGCTTCTAGCATAGCC 77
 Db

QY	1918	TCGGCCACGAGCGGTGGCGAATGAGAGATAGACATGCGATGTATGAGAGAGATCGAAGAT	1977
Db	76	TCGGCCACGAGCGGTGGCGAATGAGAGATAGACATGCGATGTATGAGAGAGATCGAAGAT	18

Oy	1978	TTFCCCGAAAAAAA	1992
Db	17	TTTCCATGCTAATAATA	3

RESULT 24
BPP62255

LOCUS	744 bp	MRNA	linear	EST 30-OCT-2000
BF162255				
601770948F1	NCL_CGAP_Lu29	Mus musculus	CDNA clone	IMAGE:39903715
DEFINITION				
mrna sequence.				

ACCESSION	BF162255
VERSION	BF162255.1
KEYWORDS	GI:11042462
SOURCE	EST.
	house mouse.

REFERENCE
1 (bases 1 to 744)
NITH-MGC <http://mcg.nci.nih.gov/>
AUTHORS
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus

TABLE
JOURNAL
COMMENT
 National Institutes of Health, Mammalian Genome Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov

CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.

found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLAM9201 row: 0 column: 12
 High quality sequence stop: 730.

```

SOURCE
1. .744
/organism="Mus musculus"
/strain="CZECH II (feral)"

```

```

/clone="IMAGE:3990371"
/clone_1lb="NCI_CGAP_Lu29"
/tissue_type="spontaneous tumor, metastatic to mammary.

```

```

/lab_node="DH10B"
/ote="Organ: lung; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT
Library constructed by life Technologies. Investigator

```

BASE COUNT	ORIGIN
201 a	183 c
212 g	148 t

Query Match	24.4%	Score 489.6	DB 10	Length 744
Best Local Similarity	84.7%	Pred. No. 16e-81		
Matches 608; Conservative	0	Mismatches 104	Indels 6	Gaps 5

20 ACCGACACCCGCGCCCTCAGCAAGCCGAGCTGAGCAGTTCTTGGCTGGCTGGCC 85
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 16 ACCGCGCGCGGTTGGCCACCGAACTGTGTGTGCTG-GCAGCTGGTTCCTGGGCT 74

75 GCTTGAGTCGAGCCATCATGCTTATCCGGCGCTGTGACACTATCTGCTCCGACTTCTTCG 134

Db	195	TCCAGTGGTTTGAGACAGACCAAGTCGAGACTGCCACAGGTAGAAATCCAGGTTGGCA	254
Qy	266	AAAGAACCATTTATCATTAAGCTCTTCTTGTGATCTTGCCACAGAGAGAGATGTCTTGG	325

Db 255 AAAAGACTATTATAACAACCTTTCTTTGACCTCGCCAGGAGAGAGAAATGCTTGG 314

Oy 326 ATGAGAAATTCCTTAAAGATGACGACAAATGTCAGAGCCAGCTTCCCAAGAACACA 385

DB 315 ATGCGAAATCTTTAAAGAAATGAACTGGACACAGCGTCAAGATCCAGCTTCCACGAAAGACA 3/4

QY 386 AGGAGAAACGAGACAGCCAGTCATTCGACACCTCTGCGGATAGCGTGAAGAAGCGCA 445

446 ATGCTACTGTGGTATCTCTGCAGCAGGCCCTTGGGCAAGGCCGAGATGCTGTGCTCCACAC 505

Query Match	23.8%	Score 476.8;	DB 10;	Length 687;
Best Local Similarity	83.1%;	Pred. No. 3.9e-79;		
Matches 555; Conservative	0;	Indels 1;	Gaps 1;	

CDNA Library Preparation: Stratagene, Inc., David B. Krizman, Ph.D.

CDNA Library Arraying: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LINL at:
www.bio.liml.gov/bdrr/image/image.html
Insert Length: 1279 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 436.

FEATURES

source
1..523
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1016379"
/clone_lib="NCI_CGAP_GC2"
/tissue_type="germ cell tumor"
/lab_host="SOLR (kanamycin resistant)"
/note="Vector: Bluescript SK-; Site_1: EcoRI; Site_2: XhoI
; Cloned unidirectionally. Primer: Oligo dt. Bulk germ
cell tumor. 5' adaptor sequence: 5' GAATTCGGCAGAG 3' 3'
adaptor sequence: 5' CTCGAGTTTCTTTTCTTTTCTTTT 3' Average
insert size: 1.2 kb."
BASE COUNT 117 a 149 c 133 g 124 t
ORIGIN

Query Match 23.1%; Score 462.8; DB 9; Length 523;
Best Local Similarity 97.7%; Pred. No. 1,7e-76;

Matches 512; Conservative 0; Mismatches 7; Indels 5; Gaps 4;

QY 1488 GCTGGACACCTTCTGTCGTGAGAAAGAGTGTGACCAATGGCCAGACATGCC 1547
|||||
DB 523 GCTGGACACCTTCTGTCGTGAGAAAGAGTGTGACCAATGGCCAGACATGCC 464
|||||
QY 1548 TGCACATTTGATGATGACAGAGTGTGACAGAGG--TTTGTGACAGACCTTCTTTCG 1605
|||||
DB 463 TGCACATTTGATGATGACAGAGTGTGACAGAGGTTTGTGACAGAGCCCTTTCG 404
|||||
QY 1606 GGACACAGCTGAGGTGTAAGGAGCAGACAAAGAGTGAAGGTGAGTGTGACACCCAGAGAC 1665
|||||
DB 403 GGACACAGCTGAGGTGTAAGGAGCAGACAAAGAGTGAAGGTGAGTGTGACACCCAGAGAC 344
|||||
QY 1666 TGCATCTTCGCTCCCTCAACCTGCCCCCACTCTACAGCTGGAGAGTGAATGACAGCCCA 1725
|||||
DB 343 TGCATCTTCGCTCCCTCAACCTGCCCCCACTCTACAGCTGGAGAGTGAATGACAGCCCA 284
|||||
QY 1726 CTGATCTCTGACAGAGTCTGCT--CTGTGACAGAGCTCTTGTATTATGACCATATCAGA 1784
|||||
DB 283 CTGATCTCTGACAGAGTCTGCT--CTGTGACAGAGCTCTTGTATTATGACCATATCAGA 224
|||||
QY 1785 TGTGTTGAGACTTTTGTGGGCTGAGAGACCAAGTCACTGTTGACATGCTCTGTGGAC 1844
|||||
DB 223 TGTGTTGAGACTTTTGTGGGCTGAGAGACCAAGTCACTGTTGACATGCTCTGTGGAC 164
|||||
QY 1845 CAGAGTCTGAGGATCTCAGGAGAGCTCAGCCCAAGCTTCTACCTGCTTGTACTTGC 1904
|||||
DB 163 CAGAGTCTGAGGATCTCAGGAGAGCTCAGCCCAAGCTTCTACCTGCTTGTACTTGC 104
|||||
QY 1905 TTCTT-AGCATAGCTGGGCGCAGAGAGGTGGGGAATGAGAGATGACATGGGATGTATGG 1963
|||||
DB 103 TTCTT-AGCATAGCTGGGCGCAGAGAGGTGGGGAATGAGAGATGAG-CATGGAGATGTATGG 45
|||||
QY 1964 AGAGAGTGAAGATTTTCCGAAAAAAGAAAAAAGAAAAA 2007
|||||
DB 44 AGAGAGTGAAGATTTTCCGAAAAAAGAAAAAAGAAAAA 1

RESULT 27

BE031265 547 bp mRNA linear EST 09-JUL-2000
LOCUS BE031265 1297770 MARC 1P16 Sus scrofa CDNA 5', mRNA sequence.
ACCESSION BE031265
VERSION BE031265.1 GI:8326274
KEYWORDS EST.

SOURCE

ORGANISM pig.

Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.

REFERENCE 1 (bases 1 to 547)
Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E.,
Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Laegreid,W.W.
and Keele,J.W.

AUTHORS Design and use of two pooled tissue normalized cDNA libraries for
EST discovery in swine

TITLE Unpublished (2000)

JOURNAL Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390

COMMENT Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt-trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -mismatch 12 options.

PCR Primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTCCTCAGTCACGACG
Plate: 65 row: H column: 4
Seq primer: ATTAGTGTGACACTATG.
Location/Qualifiers
1..547
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone_lib="MARC 1P16"
/tissue_type="pooled"
/lab_host="DH10B"
/note="Vector: PCMV SpORF6; Site_1: XbaI; Site_2: XhoI;
Library made from pooled tissue from day 11, 13, 15, 20,
and 30 embryos."

FEATURES

source
BASE COUNT 166 a 130 c 161 g 90 t
ORIGIN

Query Match 23.0%; Score 461.2; DB 9; Length 547;
Best Local Similarity 90.3%; Pred. No. 3.3e-76;

Matches 493; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 210 GTCTTTGAGACACACAGTGTGACCTGCGCCAGAGCGGAATCCAGTTGGCAAAAG 269
|||||
DB 2 GTCTTTGAGACACACAGTGTGACCTGCGCCAGAGCGGAATCCAGTTGGCAAAAG 61
|||||
QY 270 AACCATTAATCAATTAAGCTCTTCTTGTGATCTTGTGACAGAGAGAGAGATGCTTGATCG 329
|||||
DB 62 AACCATTAATCAATTAAGCTCTTCTTGTGATCTTGTGACAGAGAGAGAGATGCTTGATCG 121
|||||
QY 330 AGAATTTCTTAAGATGAATGACATGTCAGAGCCAGCTTTCCAGAAAGCAAGGA 389
|||||
DB 122 AGAATTTCTTAAGATGAATGACATGTCAGAGCCAGCTTTCCAGAAAGCAAGGA 181
|||||
QY 390 GAAACGAGACAGCGAGGATCATCGACACTCTCGGGATTAACCTGGAAGAACCAATGCG 449
|||||
DB 182 GAAACGAGACAGCGAGGATCATCGACACTCTCGGGATTAACCTGGAAGAACCAATGCG 241
|||||
QY 450 TACTGTGATCTCTGACAGAGGCTTGGCAAGGCGAGATGCTGTGCTCCACATGAA 509
|||||
DB 242 CACTGTGGAATCCCTCGAAGAGGCTTGAACAAAGCGGAATGCTGTGCTCCACATGAA 301
|||||
QY 510 AAACGAGATGAATGATTAAGCAGCAGCAGAGATGAGAACCAACAGCAGAGAGGCG 569
|||||
DB 302 AAACGAGATGAATGATTAAGCAGCAGCAGAGATGAGAACCAACAGCAGAGAGGCG 361
|||||
QY 570 GGCCGGCTGAGAGCAAGATGAAGACCATGAGATGAGTGTGCTTACTCCAGAGCA 629
|||||
DB 362 CGGCCGACTCCGAGCAGATGAGAGTATGAGAGCGATTGAGTCTTACTCCAGAGCA 421
|||||
QY 630 GCTCCCTGAGGTGAGAGATGATCGAGACATGGGTGTGGACAGTCCAGCGGTGAGACA 689
|||||

Db 422 GCGCGCTGAGGTGAGAAATGATCCGAGAAATGGGTGTGGACATCAGCAGTGGACGA 481
 QY 690 GCTGCTCTGACTGTGTCTCTCTCAAGAAAGATGAGATCTAAAGAGGACGAGAA 749
 Db 482 GCTGCTCTGACTGTGTCTCTCTCAAGAAAGATGAGATCTAAAGAGGACGAGAA 541
 QY 750 GGCCTC 755
 Db 542 GGCCTC 547
 RESULT 28
 B1341090 571 bp mRNA linear EST 30-JUL-2001
 LOCUS 368355 MARC 2Pig Sus scrofa cDNA 5', mRNA sequence.
 DEFINITION B1341090
 ACCESSION B1341090.1 GI:15034379
 VERSION EST.
 KEYWORDS EST.
 SOURCE Sus scrofa
 ORGANISM Sus scrofa
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
 1 (bases 1 to 571)
 Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E.,
 Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Laegreid,W.W.
 and Keele,J.W.
 Design and use of two pooled tissue normalized cDNA libraries for
 EST discovery in swine
 Unpublished (2000)
 JOURNAL Contact: Smith TPL
 USDA, ARS, US Meat Animal Research Center
 PO Box 166, Clay Center, NE 68933-0166, USA
 Tel: 402 762 4366
 Fax: 402 762 4390
 Email: smithemail.marc.usda.gov
 Single pass sequencing. Bases called and alt-trimmed with phred
 v0.980904.e. Vector identified by cross_match with the -minscore 18
 and -mismatch 12 options.
 PCR primers
 FORWARD: AGCAACAGCTATGACCAT
 BACKWARD: GTTTCCTCAGTCACGACG
 Plate: 107 row: K column: 10
 Seq primer: ATTTAGCTGACACTATAG.
 Location/Qualifiers
 1..571
 /organism="Sus scrofa"
 /db_xref="taxon:9823"
 /clone_id="MARC 2Pig"
 /rissue_type="pooled"
 /lab_host="DH10B"
 /note="Vector: pCMV SPORT6; Site.1: XbaI; Site.2: XhoI;
 Library made from pooled tissue from testis, ovary,
 endometrium, hypothalamus, pituitary, and placenta."
 BASE COUNT 154 a 162 c 146 g 109 t
 ORIGIN
 Query Match 22.1%; Score 443; DB 10; Length 571;
 Best Local Similarity 88.3%; Pred. No. 8,1e-73;
 Matches 506; Conservative 0; Mismatches 60; Indels 7; Gaps 2;

QY 187 TTCCACTTGCAGTGCCTATTCAGTCTTTGAGACAGCACCACCAAGTCGACCTGCCACAG 246
 Db 178 TTCCACTTGCAGTGCCTATTCAGTCTTTGAGACAGCACCACCAAGTCGACCTGCCACAG 237
 QY 247 TGCCGAATCCAGGTTGGCAAAAGAACCTATATCAATTAAGCTCTTTGATCTGGCCAG 306
 Db 238 TGCCGAATCCAGGTTGGCAAAAGAACCTATATCAATTAAGCTCTTTGATCTGGCCAG 297
 QY 307 GAGGAGGAGATGTCTTGATCGAATTCCTTAAGAAATGAACAGCAATGTCAGGCC 366
 Db 298 GAGGAGGAGATGTCTTGATCGAATTCCTTAAGAAATGAACAGCAATGTCAGGCC 357
 QY 367 CAGCTTTCCAGAAAGACAG 426
 Db 358 CTGCTTTCCAGAAAG 417
 QY 427 GATACGCTGGAAG 486
 Db 418 GACACGCTGGAAG 477
 QY 487 GAGATGCTGTGCTCCACACCTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 546
 Db 478 GAAATGCTGTCTCCACACCTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 537
 QY 547 ACCAAACAGCACAAG 579
 Db 538 ACCAAACAGCACAAG 570

RESULT 29
 BG235999/c 495 bp mRNA linear EST 12-FEB-2001
 LOCUS naf20f09.x1 Soares_NPBMC Homo sapiens cDNA clone IMAGE:4141600 3',
 DEFINITION mRNA sequence.
 ACCESSION BG235999
 VERSION BG235999.1 GI:12749846
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 495)
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 JOURNAL Contact: Robert Strausberg, Ph.D.
 COMMENT Email: cgaps-remail.nih.gov
 Tissue Procurement: M. Bento Soares, Ph.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D. and M.Fatima
 Bonaldo, Ph.D.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
infoimage.lnl.gov
 Seq primer: ~400p from Gibco
 High quality sequence stop: 416.
 Location/Qualifiers
 1..495
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_id="IMAGE:4141600"
 /clone_lib="Soares_NPBMC"
 /rissue_type="lymphocyte"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: blood; Vector: pT73D-Pac; Site.1: NotI;
 Site.2: EcoRI; 1st strand cDNA was primed with a Not I -
 oligo(dT) primer [5'
 TGTTACCAATCTGACAGTGGAGCGCCGCGGCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT
 3']; double-stranded cDNA was ligated to Eco RI adaptors
 (pharmacia), digested with Not I and cloned into the Not
 I and Eco RI sites of the modified pT73 vector. Library

is normalized; constructed in the Laboratory of M. Bento Soares (University of Iowa).
 BASE COUNT 111 a 137 c 123 g 124 t
 ORIGIN

Query Match 21.1%; Score 422.6; DB 10; Length 495;
 Best Local Similarity 97.1%; Pred. No. 5.2e-69;
 Matches 473; Conservative 0; Mismatches 9; Indels 5; Gaps 4;

QY 1525 TGACCAATGGCCAGACACATGCTTGCACACTTGTAGTCAAGGACTGTCAGGCGAGG--T 1562
 DB 495 TGACCAATGGCCAGACACATGCTTGCACACTTGTAGTCAAGGACTGTCAGGCGAGGTT 436
 QY 1583 TTGTGACACAGCCCTTCTTGGGACACAGCTGAGGTGAAGGCGACACAACAGGTGA 1642
 DB 435 TTGTGACACAGCCCTTCTTGGGACACAGCTGAGGTGAAGGCGACACAACAGGTGA 376
 QY 1643 GGGTGAAGTGTGACACCCAGACAGTGTCTTCTGCCCCACCTGCCCCACCTCTTACGAC 1702
 DB 375 GGGTGAAGTGTGACACCCAGACAGTGTCTTCTGCCCCACCTGCCCCACCTCTTACGAC 316
 QY 1703 TGGAGCTGCATGACACAGCCACCTATCTGTGACAGAGTCTGCT-CTGTTGCCAGGC 1761
 DB 315 TGGAGCTGCATGACACAGCCACCTATCTGTGACAGAGTCTGCTGTTGCCAGGC 256
 QY 1762 TCTTGTATAGGCATGATCAGATGTGTGACACTTCTTGGGCGCTGGAGACAGGTC 1821
 DB 255 TCTTGTATAGGCATGATCAGATGTGTGACACTTCTTGGGCGCTGGAGACAGGTC 196
 QY 1822 ACTTGTGACTGTCTGTGTGACAGAGTCTTGTGAGCATCTGACGACGCTTACGCCAA 1881
 DB 195 ACTTGTGACTGTCTGTGTGACAGAGTCTTGTGAGCATCTGACGACGCTTACGCCAA 136
 QY 1882 GCTTTACTGCTCTTGTACTTGTCTTA-GCATAGCTTGGGCCAAGCAGGTTGGGATG 1940
 DB 135 GCTTCTACCTGCTTGTACTTGTCTTGTAGCATAGCTTGGGCCAAGCAGGTTGGGATG 76
 QY 1941 GAGATGACATGAGATGTATGAGAGATGGAAGATTTCCGAAAAA 2000
 DB 75 GAGATGAG-CATGGAGATGTATGAGAGATGGAAGATTTTCAATAAATTAAAA 17
 QY 2001 AAAAAA 2007
 DB 16 AAAAAA 10

RESULT 30
 AM248701 478 bp mRNA linear EST 07-JAN-2000
 LOCUS 2821007.3pr1me NIH_MGC_7 Homo sapiens cDNA clone IMAGE:2821007 3',
 DEFINITION mRNA sequence.
 ACCESSION AM248701
 VERSION AM248701.1 GI:6591694
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 478)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Other_ESTS: 2821007.5pr1me
 Contact: Robert Strausberg, Ph.D.
 Email: cgaps-femail.nih.gov
 Tissue Procurement: DCTD/DTP cDNA Library Preparation: Ling
 Hong/Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E.
 Consortium (LNL) DNA Sequencing by: Berkeley MGC sequencing
 project Clone distribution: MGC clone distribution information can
 be found through the I.M.A.G.E. Consortium/LNL at:
www-bio.lnl.gov/bdrip/image/image.html Base Calling / Quality
 Scores: PHRED from University of Washington Genome Center. Vector

Trimming: cross-match from University of Washington Genome Center
 PHRAP suite. Poly-T identification: patmatch.pl from Berkeley
 Drosophila Genome Project. Universtion of Washington Genome Center:
<http://www.genome.washington.edu> polydenylation: Based upon the
 presence of a XhoI site followed by a run of 14 or more T residues
 at the beginning of the sequence, this cDNA insert was
 polyadenylated.
 Plate: LICMS row: K column: 24
 High quality sequence stop: 458.

FEATURES
 source
 location/Qualifiers

1..478
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2821007"
 /clone_1ib="NIH_MGC_7"
 /tissue_type="small cell carcinoma"
 /cell_line="MGC3"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: lung; Vector: pOT87; Site_1: XhoI; Site_2:
 EcoRI; cDNA made by oligo-dT priming. Directionally
 cloned into EcoRI/XhoI sites using the following 5'
 adaptor: GGCAAGAG(G). Size selected >500bp for average
 insert size 1.8kb. Library constructed by Ling Hong in
 the Laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies)."
 BASE COUNT 111 a 134 c 120 g 112 t 1 others
 ORIGIN

Query Match 20.4%; Score 410.4; DB 9; Length 478;
 Best Local Similarity 96.5%; Pred. No. 9.9e-67;
 Matches 462; Conservative 0; Mismatches 12; Indels 5; Gaps 4;

QY 1532 TGCCACAGACACATGCTTGTAGTCAAGGACTGTCAGGCGAGG--TTTGTGA 1589
 DB 478 TGCCACAGACACATGCTTGTAGTCAAGGACTGTCAGGCGAGGTTTGTGA 419
 QY 1590 CAGAGCCCTACTTGTGGGACACAGCTGAGGTGAAGGCGACACAAGAGTGAAGGTGAG 1649
 DB 418 CAGAGCCCTACTTGTGGGACACAGCTGAGGTGAAGGCGACACAAGAGTGAAGGTGAG 359
 QY 1650 TGTGACACCCAGAGACTGTCTTCTGCTCCCTCAACCTGCCCCACTCTTACGACTGGAGC 1709
 DB 358 TGTGACACCCAGAGACTGTCTTCTGCTCCCTCAACCTGCCCCACTCTTACGACTGGAGC 299
 QY 1710 TGACATGACACAGCCACTGATCTGTGACAGAGTCTGCT-CTGTTGCCAGGCTCTTGT 1768
 DB 298 TGACATGACACAGCCACTGATCTGTGACAGAGTCTGCTCTGTTGCCAGGCTCTGTT 239
 QY 1769 TATAGCCATGATGATGTGTGTGACAGTCTTGTGAGGCTTGGAGACACAGGTCATTGTT 1828
 DB 238 TATAGCCATGATGATGATGTGTGTGACAGTCTTGTGAGGCTTGGAGACACAGGTCATTGTT 179
 QY 1829 GACTGTCTGTGTGACACAGAGTGTGAGCATCTGACGACGCTTCAAGCCCAAGCTTCTA 1888
 DB 178 GACTGTCTGTGTGACACAGAGTGTGAGCATCTGACGACGCTTCAAGCCCAAGCTTCTA 119
 QY 1889 CTTGCTTTTGAATGCTTCTTA-GCATAGCCTGGGCCAAGCAGGTTGGGATGAGAGATA 1947
 DB 118 CTTGCTTTTGAATGCTTCTTGTAGCATAGCCTGGGCCAAGCAGGTTGGGATGAGAGATA 59
 QY 1948 GACATGGAGATGTAGAGAGATGGAAGATTTTCCGAAAAA 2006
 DB 58 G-CATGGAGATGTAGAGAGATGGAAGATTTTCAATAAATAAATAAATAA 1

RESULT 31
 BE304189 519 bp mRNA linear EST 13-JUL-2000
 LOCUS BE304189
 DEFINITION 601086421F1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:3500621 5',
 mRNA sequence.
 ACCESSION BE304189

VERSION	BE304189.1	GI:9174333
KEYWORDS	EST.	
SOURCE	house mouse.	
ORGANISM	Mus musculus	
REFERENCE	<p>NIH-MGC http://mgc.ncl.nih.gov/.</p> <p>1 (bases 1 to 519)</p> <p>Mammalia: Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.</p>	
AUTHORS	National Institutes of Health, Mammalian Gene Collection (MGC)	
TITLE	Unpublished (1999)	
JOURNAL	Contact: Robert Strausberg, Ph.D.	
COMMENT	<p>Email: cgapbs-rr@mail.nih.gov</p> <p>Tissue Procurement: Jeffrey Green M.D.</p> <p>cDNA Library Preparation: Life Technologies, Inc.</p> <p>CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLU)</p> <p>DNA Sequencing by: Incyte Genomics, Inc.</p> <p>Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLU at: http://lmgc.lnl.gov</p> <p>Plate: LLM8359 row: 1 column: 06</p> <p>High quality sequence stop: 515.</p>	
FEATURES	Location/Qualifiers	
	1..519	
source	<p>/organism="Mus musculus"</p> <p>/strain="FVB/N"</p> <p>/db_xref="taxon:10090"</p> <p>/clone="IMAGE:3500621"</p> <p>/clone_lib="NCI CGAP Mam6"</p> <p>/sex="female, virgin"</p> <p>/tissue_type="infiltrating ductal carcinoma"</p> <p>/dev_stage="5 months"</p> <p>/lab_host="DH10B"</p> <p>/note="Organ: mammary; Vector: PCMV-SPOrt6; Site_1: SalI; Site_2: NciI; Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies. Investigator providing samples: Jeffrey Green, M.D., NIH"</p>	
BASE COUNT	159 a	116 c 152 g 92 t
ORIGIN		
Query Match	20.2%	Score 406.2; DB 10; Length 519;
Best Local Similarity	86.8%;	Pred. No. 5.9e-66;
Matches 447; Conservative	0; Mismatches 68; Indels	0; Gaps 0
OY	254	TCCAGGTGGCAAAAGAACCATTTATCAATTAACCTCTCTTTTATCTTGGCCAGAGAGG 313
DB	5	TCCAGGTGGCAAAAGAACTATTATTAACAACCTTTCTTGACCTGCCAGAAAGAG 64
OY	314	AGAAATGCTTGGATCGAGCAATTCCTTAAGAAATGAAGTGGACAAATGTCAGAGCCCACTTT 373
DB	65	AGAAATGCTTGGATCGAGCAATTCCTTAAGAAATGAAGTGGACAGCGCTCAAGCTCAGCTTT 124
OY	374	CCCAAGAGACAAGAGAAAGACAGACAGCCAGCTCATCATCGACACTCTGGCGGATAGCC 433
DB	125	CCCAAGAGAGACAGGAGAAAGCGGACAGCCAGCATTTATGACACTCTAGCGGAGACCC 184
OY	434	TGGAAGAAAGCAATCTACTTGTGATCTCTGACAGAGCCTTTGGGCAAGCGGATAGC 493
DB	185	TGGAAGAAAGCAATCTCTACCGTGGATGCCCTTACAGAACCCCTTAAACAAGCAGAGATGC 244
OY	494	TGTGCTCCACACTGAAAAAGACAGATGAAGTACTTGAAGCAGCAGCAGATGAGACCAAC 553
DB	245	TGTGTTCCACCTCGAAAAAAGACAGATGAAGTCTCTGGAGCAGCGCAGATGAGACCAAC 304
OY	554	AAGCAACAGAGAGAGCGGCGGCTCAGAGCAAGATGAAGACCATGAGCAGATTGAGC 613
DB	305	AAGCTCGGAGAGAGGCCCAACCGACTCAAGTGAAGTGAACCAATGGAGCAAAATTGAGC 364
OY	614	TTCCTACTCCAGAGCCAGCTCCCTGAGAGTGGAGAGATGATCCGAGACATGGGTGGGAG 673
DB	365	TCTCTACTCCAGAGCCAGCGGCTTGAGAGTGGAGAGATGATTTGAGACATGGGTGGGAG 424
OY	674	AGTCAGCGGTGAGACAGCTGAGCTGTGTACTGTGTCTCTCAAGAAAGATGACGAAATC 733

[illegible]

cdna Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.
cdna Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bdrip/image/image.html
Seq primer: -40ml3 fwd. E7 from Amersham
High quality sequence stop: 425.
Location/Qualifiers
1. 435

FEATURES
SOURCE
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1839888"
/clone_lib="Soares_testis_NHT"
/sex="male"
/lab_host="DH10B"
/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from Clontech Laboratories, Inc., and primed with a Not I - oligo(dT) primer [5', TGTTACCAATCTGACAGTGGAGCGGCCCAATTTTCTTTTCTTTT 3']. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization to Cots, and was constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT 103 a 124 c 112 g 95 t 1 others
ORIGIN

Query Match 18.6%; Score 373.2; DB 9; Length 435;
Best Local Similarity 97.0%; Pred. No. 8,7e-60;
Matches 423; Conservative 0; Mismatches 8; Indels 5; Gaps 4;
QY 1568 CTGTCCAGGAGG--TTTGTGACAGAGCCCTACTTTCGGGACCGCTGAGGTGAAG 1635
|||||
DB 435 CTGTCCAGGAGGTTTGTGTGACAGAGCCCACTTTCGGGACCGCTGAGGTGAAG 376
QY 1636 GGCAGACAAACAGGTGGAGGTGAGTGACACCCAGACAGCTGCTCCCTCACC 1665
|||||
DB 375 GGCAGACAAACAGGTGGAGGTGAGTGACACCCAGACAGCTGCTCCCTCACC 316
QY 1686 TGCCCACTCTTACGACGTGAGGTGACATACACGCCCACTGATCCTGTACAGAGTCC 1745
|||||
DB 315 TGCCCACTCTTACGACGTGAGGTGACATACACGCCCACTGATCCTGTACAGAGTCC 256
QY 1746 TGCT-CTGTGCGCAGGCTCTTTTATAGCATGATGATGTCAGACTCTTCTGG 1804
|||||
DB 255 TGCTCTGTGTCACAGGCTCTGTTTATAGCATGATGATGTCAGACTCTTCTGG 196
QY 1805 GCGTGGAGACAGGTGACCTGTTGACGTCTGTGTGACACAGATGCTTATAGCATCTC 1864
|||||
DB 195 GCGTGGAGACAGGTGACCTGTTGACGTCTGTGTGACACAGATGCTTATAGCATCTC 136
QY 1865 AGCAGGCTCAGCCCAAGCTTCTACCTGCTTGTACTTCTCT-AGCATAGCTGGGCC 1923
|||||
DB 135 AGCAGGCTCAGCCCAAGCTTCTACCTGCTTGTACTTCTCTTACGATAGCTGGGCC 76
QY 1924 AAGCAGGTGGGGAATGAGATGACATGAGATGATGAGAGAGAGGAAGATTTTCCC 1983
|||||
DB 75 AAGCAGGTGGGGAATGAGATGAGATGATGAGAGAGAGGAAGATTTTCAAT 17
QY 1984 GAAAAAATTTTAAAA 1999
|||||
DB 16 GTAAAAATTTAAAA 1

RESULT 35
AM207517/c 455 bp mRNA linear EST 02-DEC-1999
LOCUS
DEFINITION UT-H-B11-afj-f-12-0-UT.s1 NCI_CGAP_Sub3 Homo sapiens cDNA clone
IMAGE:2721959 3', mRNA sequence.

ACCESSION AM207517
VERSION AM207517.1 GI:6507013
KEYWORDS EST
SOURCE human
ORGANISM Homo sapiens
REFERENCE Eukaryota: Metazoa: Chordata: Cranialata: Vertebrata: Euteleostomi; Mammalia: Eutheria: Primates: Catarrhini: Homnidae: Homo. 1 (bases 1 to 455)
NCT-CGAP <http://www.ncbi.nlm.nih.gov/nctgap>.
National Cancer Institute, Cancer genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapb-remail.nih.gov
The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. cDNA Library Preparation: M.B. Soares Lab Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bdrip/image/image.html The following repetitive elements were found in this cDNA sequence: 1-29,
>AT-rich#Low-complexity
Seq primer: M13 Forward
POLYA=yes.

FEATURES
SOURCE
Location/Qualifiers
1. 455
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2721959"
/clone_lib="NCI_CGAP_Sub3"
/lab_host="DH10B (Life Technologies)"
/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; The NCI_CGAP_Sub3 library is a subtracted library derived from the NCI_CGAP_Sub1 library, which is a subtracted library derived from BI. BI constitutes a mixture of 21 normalized or subtracted NCI_CGAP libraries: NCI_CGAP_C04, NCI_CGAP_Pr22, NCI_CGAP_Pr28, NCI_CGAP_C010, NCI_CGAP_C016, NCI_CGAP_Kid5, NCI_CGAP_Kid12, NCI_CGAP_Kid3, NCI_CGAP_Kid11, NCI_CGAP_Lym2, NCI_CGAP_Br2, NCI_CGAP_C08, NCI_CGAP_C0L1, NCI_CGAP_L0412, NCI_CGAP_Brn23, NCI_CGAP_Lu5, NCI_CGAP_Lu24, NCI_CGAP_Lu19, NCI_CGAP_G04, NCI_CGAP_G06, NCI_CGAP_Brn25. These 21 libraries were pooled and a single-stranded DNA preparation of the resulting mixture was used as a tracer in a subtractive hybridization with a driver whose composition is detailed below:
NCI_CGAP_Kid3 pool 1 LLAM 3334-3337, 3682-3683,
3798-3803 (IMAGE Clonoids 1322376-1323911, 1456008-1456775, 1500552-1502855); NCI_CGAP_Kid5 pool 1 LLAM 3338-3342, 3722-3725, 3776-3778 (IMAGE Clonoids 1322912-1325831, 1471368-1472803, 1492104-1493253); NCI_CGAP_Lu5 pool 1 LLAM 3575-3582, 3851-3854 (IMAGE Clonoids 1414920-1417991, 1520904-1522439); NCI_CGAP_G04 pool 1 LLAM 3164-3167, 3716-3720, 3733-3735 (IMAGE Clonoids 1257096-1258631, 1469064-1470983, 1475562-1476433); NCI_CGAP_Pr22 pool 1 LLAM 2457-2459, 2758-2759, 3062-3068 (IMAGE Clonoids 985608-986759, 1101192-1101959, 121928-1220615); NCI_CGAP_C010 pool 1 LLAM 2644-2653, 2871-2872 (IMAGE Clonoids 1057416-1061255, 1144584-1145351). Subtraction was performed as previously described (Bonaldo, Lennon & Soares (1996): Normalization and subtraction: Two Approaches To Facilitate Gene Discovery. Genome Research 6, 791-806.
TAG_LIB=NCI_CGAP_C04
TAG_TISSUE=colon
TAG_SEQ=CTCG"

BASE COUNT 103 a 129 c 113 g 108 t 2 others
ORIGIN
Query Match 18.2%; Score 365; DB 9; Length 455;

PCR Primers
FORWARD: AGGAACACGCTATGACCAT
BACKWARD: GTTTCCCGACGCACGACG
Plate: 77 row: 0 column: 14
Seq primer: ATTTAGGTGACACTATAG

BASE COUNT	118 a	130 c	115 g	94 t
ORIGIN				

Query Match	17.28;	Score 346;	DB 9;	Length 457;
Best Local Similarity	87.68;	Pred. No. 1e-54;		
Matches 403; Conservative	0;	Mismatches 50;	Indels 7;	Gaps 2

OY	10	AGCGAAATTTGAAGCAAGGGGAGG	---	GGGCGCTCTAGCAAGCGGCAACCTGTAGAC	65
Db	1	AGCGAAATTTGAAGCAAGCTGGAGCGCTGAGGACAGCGCCACTTAAGCTGGAACCTTCGGCG			60
OY	66	TTTCTTTGGCTGCCTGGGCCCCCTTGAGTCCAGCCATCATGCTTATCCGTCTGTGCAC			125
Db	61	GTTCCTTGGTACCTGGGCACTTAGACACAG	---	TCATGGCTATTGGTGTGTGGAC	117
OY	126	TATCTGCTCCGACTTCTTCGATCATCTCCCGGACGTGGCGGCATTCATCTGGGGCACAC			185
Db	118	TATATGCTCCGACTTCTTCGATCATCTCCCGCAGCTGGGACCATTCATCTGGGGCACAC			177
OY	186	CTTCCACTTGGAGTGCCTTAATTCACTGCTTTGAGACAGACCAAGTCGACCTGGCCACA			245
Db	178	CTTCCACTTGGAGTGCCTTAATTCACTGAGGGGTGCAAGACAGACCAAGTCGACCTGGCCACA			237
OY	246	GTCGCCAATCCAGGTGGGCAAAAAGCAACCTTATTCATAAATGACTCTTCTTTGATCTTGGCCA			305
Db	238	GTCGCCAATCCAGGTGGGCAAAAAGCAACCTTATTCATAAATGACTCTTCTTTGACCTTGGCCA			297
OY	306	GGAGAGAGAGATGCTTGGATCGAGAATTCCTTAAGAATGAACGTGCAATGTCAAGC			365
Db	298	GGAGAGAGAGATGCTTGGATCGAGAATTCCTTAAGAATGAACGTGATGAATTAACAGAAC			357
OY	366	CCAGCTTTCAGAAAGCAAGAGAGAAAAGACACAGCCAGGTCAATCATCGACACTCTGCG			425
Db	358	CCCTGCTTCCAGAAAGAGAAAGAAAAGACAGACAGCCAGATCATTCATTGACACATTGGG			417
OY	426	GGATACGCTGGAAGAACCGCAATGCTACTGTGGTATCTCTG			465
Db	418	GGACACGCTGGAAGAGCGCAACCGCACATCTGGAATTCCTG			457

RESULT	38
BM462511	
LOCUS	BM462511
DEFINITION	1015 bp mRNA
	AGENCOURT_6426349 NIH_MGC_71 Homo sapiens CDNA clone IMAGE:518193
	5', mRNA sequence.

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (pages 1 to 1015)	NIH-MGC	http://mgc.ncl.nih.gov/	National Institutes of Health, Mammalian Gene Collection (MGC)	Unpublished (1999)
Contact:	Robert Strausberg, Ph.D.			

Email: cgarabs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: L1AM12178 row: j column: 18
High quality sequence stop: 644.

BASE COUNT	251 a	278 c	278 g	207 t	1 others
ORIGIN					

	Query Match	17.2%	Score 345.4:	DB 10:	Length 1015:
	Best Local Similarity	96.1%:	Pred. No. 1.2e-54:		
	Matches 366:	Conservative 0:	Mismatches 11:	Indels 4:	Gaps 1:
QY	6 GTGAGCGCAAAATTGTGAACAAAGCGGAGCGGGGCGC---CTCTAGAAAGCCGGACTGTGA 61				
Db	8 GTCCGGGCCAAATTTGAACAAGCGGAGCGGGGGCGGTCGTACGAAAGCCGGACTGTGA 67				
QY	62 GCATTTTCTTTGGCTGCCTGGGCCCCCTTGATGTCACGCCATCATGCTTATCGCTGCTGT 121				
Db	68 GCAATTTCTTTGGCTGCCTGGGCCCCCTTGATGTCACGCCATCATGCTTATCGCTGCTGT 127				
QY	122 GCATATCTGCTCCGACTCTTTCGATCACTCCCGAGAGTGGCGGCATCACTGCGGCC 181				
Db	128 GCATATCTGCTCCGACTCTTTCGATCACTCCCGAGAGTGGCGGCATCACTGCGGCC 187				
QY	182 ACACCTTTCACCTTGCAGTGCCTAATTCAGTCCCTTTGAGACACACCAAGTGGACTGCG 241				
Db	188 ACACCTTTCACCTTGCAGTGCCTAATTCAGTCCCTTTGAGACACCAAGTGGACTGCG 247				
QY	242 CACAGTGGCGAATCCAGTGTGGCAAAAAGACATATGATAATTAAGCTCTTCTTGATCTTG 301				
Db	248 CACAGTGGCGAATCCAGTGTGGCAAAAAGACATATGATAATTAAGCTCTTCTTGATCTTG 307				
QY	302 CCCAGAGGAGAGAGAATGCTTGGATCGAATTTCTTAAGAAATGAACTGGACAAATGTCA 361				
Db	308 CCCAGAGGAGAGAGAATGCTTGGATCGAATTTCTTAAGAAATGAACTGGACAAATGTCA 367				
QY	362 GAGCCCAAGCTTCCCAGAAAG 382				
Db	368 GAGCCCAAGCTTCCCAGAAAG 388				

RESULT	39
AI633797/c	
LOCUS	AI633797
DEFINITION	tt28f05.x1 NCI_CGAP_GC6 Homo sapiens cDNA clone IMAGE:2242113 3', mRNA sequence.

VERSION	AI633797.1	GI:4685127
KEYWORDS	EST.	
SOURCE	human.	
ORGANISM	Homo sapiens	

REFERENCE
AUTHORS
TITLE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 379)

NCI-CCGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

JOURNAL
COMMENT

Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/dbirp/image/image.html
Insert Length: 314 Std Error: 0.00
Seq primer: -40UP from Gibco.

FEATURES
source

```
1..379
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2242113"
/clone.lib="NCI CGAP GC6"
/tissue.type="pooled germ cell tumors"
/lab_host="DH10B"
/note="Vector: pTR73D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; Plasmid DNA
from the normalized library NCI CGAP GC4 was prepared, and
ss circles were made in vitro. Following HAP purification,
this DNA was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from a pool
of 5,000 clones made from the same library (cloneids
1257096-1258631, 1469064-1470983, and 1475592-1476743).
Subtraction by Bento Soares and M. Fatima Bonaldo."
```

BASE COUNT
ORIGIN

```
92 a 107 c 97 g 83 t
```

Query Match 16.6%; Score 334; DB 9; Length 379;
Best Local Similarity 97.9%; Pred. No. 1.8e-52;
Matches 370; Conservative 0; Mismatches 5; Indels 3; Gaps 3;

```
QY 1619 GTGTAAAGGCGACAAACAGAGTGAAGGTGAGTGACACCCAGAGACTGCTTTCTGCC 1678
Db 379 GTGTAAAGGCGACAAACAGAGTGAAGGTGAGTGACACCCAGAGACTGCTTTCTGCC 320
QY 1679 CTCACCCCTGCCCACTCTCTGAGACTGGAGCTGACATGACAGCCCACTGATCCTGTCAG 1738
Db 319 CTCACCCCTGCCCACTCTCTGAGACTGGAGCTGACATGACAGCCCACTGATCCTGTCAG 260
QY 1739 CAGGTCTCTGCT-CTGTTGCCAGGCTCTTGTATTAGCCATGATCAGATGTCAGACTC 1797
Db 259 CAGGTCTCTGCTCTGTTGCCAGGCTCTGTTATTAGCCATGATCAGATGTCAGACTC 200
QY 1798 TTTTCTGGGCGCTGGAGACAGGTCACCTTGTGACTGTCCTGTGGACCAAGTCTTGAG 1857
Db 199 TTTTCTGGGCGCTGGAGACAGGTCACCTTGTGACTGTCCTGTGGACCAAGTCTTGAG 140
QY 1858 GCATCTCAGGACAGCCTCAGGCCAAGCTTCTACCTGCTTGGACTGGTTGTA-GCATAGC 1916
Db 139 GCATCTCAGGACAGCCTCAGGCCAAGCTTCTACCTGCTTGGACTGGTTGTA-GCATAGC 80
QY 1917 CTGGGCCCAAGCAGGCTGGGGAATGAGAGATAGACATGGATGATGAGAGAGATGAGAGA 1976
Db 79 CTGGGCCCAAGCAGGCTGGGGAATGAGAGATAG-CAATGGATGATGAGAGAGATGAGAGA 21
QY 1977 TTTTCCGAAAAAAA 1994
Db 20 TTTTCATGATAAATAAAA 3
```

RESULT 40
LOCUS AM002888/c 379 bp mRNA linear EST 08-MAR-2000
DEFINITION wt02c11.x1 NCI_CGAP_GC6 Homo sapiens cDNA clone IMAGE:2480372 3',

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

mRNA sequence.
AM002888
AM002888.1 GI:5849804
EST.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 379)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/dbirp/image/image.html
Insert Length: 482 Std Error: 0.00
Seq primer: -40UP from Gibco.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENTFEATURES
source

```
1..379
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2480372"
/clone.lib="NCI CGAP GC6"
/tissue.type="pooled germ cell tumors"
/lab_host="DH10B"
/note="Vector: pTR73D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; Plasmid DNA
from the normalized library NCI CGAP GC4 was prepared, and
ss circles were made in vitro. Following HAP purification,
this DNA was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from a pool
of 5,000 clones made from the same library (cloneids
1257096-1258631, 1469064-1470983, and 1475592-1476743).
Subtraction by Bento Soares and M. Fatima Bonaldo."
```

BASE COUNT
ORIGIN

```
92 a 107 c 97 g 83 t
```

Query Match 16.6%; Score 334; DB 9; Length 379;
Best Local Similarity 97.9%; Pred. No. 1.8e-52;
Matches 370; Conservative 0; Mismatches 5; Indels 3; Gaps 3;

```
QY 1619 GTGTAAAGGCGACAAACAGAGTGAAGGTGAGTGACACCCAGAGACTGCTTTCTGCC 1678
Db 379 GTGTAAAGGCGACAAACAGAGTGAAGGTGAGTGACACCCAGAGACTGCTTTCTGCC 320
QY 1679 CTCACCCCTGCCCACTCTCTGAGACTGGAGCTGACATGACAGCCCACTGATCCTGTCAG 1738
Db 319 CTCACCCCTGCCCACTCTCTGAGACTGGAGCTGACATGACAGCCCACTGATCCTGTCAG 260
QY 1739 CAGGTCTCTGCT-CTGTTGCCAGGCTCTTGTATTAGCCATGATCAGATGTCAGACTC 1797
Db 259 CAGGTCTCTGCTCTGTTGCCAGGCTCTGTTATTAGCCATGATCAGATGTCAGACTC 200
QY 1798 TTTTCTGGGCGCTGGAGACAGGTCACCTTGTGACTGTCCTGTGGACCAAGTCTTGAG 1857
Db 199 TTTTCTGGGCGCTGGAGACAGGTCACCTTGTGACTGTCCTGTGGACCAAGTCTTGAG 140
QY 1858 GCATCTCAGGACAGCCTCAGGCCAAGCTTCTACCTGCTTGGACTGGTTGTA-GCATAGC 1916
Db 139 GCATCTCAGGACAGCCTCAGGCCAAGCTTCTACCTGCTTGGACTGGTTGTA-GCATAGC 80
QY 1917 CTGGGCCCAAGCAGGCTGGGGAATGAGAGATAGACATGGATGATGAGAGAGATGAGAGA 1976
Db 79 CTGGGCCCAAGCAGGCTGGGGAATGAGAGATAG-CAATGGATGATGAGAGAGATGAGAGA 21
```

Db 79 CTGGGCCCAAGCAGGGGTGGGAATGAGGATAG-CATGGATGTAATGAGAGGATGAAGA 21
Qy 1977 TTTTCCCGAATAAAAAA 1994
||||| | |||| ||||
Db 20 TTTTCATCTAATAATTA 3

Search completed: September 4, 2002, 11:29:22
Job time: 10277 sec